HA
1
S764
vol. 1
1959-60
n/c

INTERNATIONAL JOURNAL OF ABSTRACTS

STATISTICAL THEORY AND METHOD ABSTRACTS

VOLUME 1 · NUMBER 1
JULY 1959 -60

PUBLISHED FOR
THE INTERNATIONAL STATISTICAL INSTITUTE
OLIVER AND BOYD

VOLUME 1 · NUMBER 1

COPYRIGHT © 1959 INTERNATIONAL STATISTICAL INSTITUTE

PRINTED IN GREAT BRITAIN BY
OLIVER AND BOYD LTD., EDINBURGH

INTERNATIONAL JOURNAL OF ABSTRACTS
STATISTICAL THEORY AND METHOD

COVERAGE OF JOURNAL

THE aim of this journal of abstracts is to give complete coverage of papers in the field of statistical theory and new contributions to statistical method. Papers which only report applications or examples of existing statistical theory and method will not be included. There are approximately two hundred and fifty journals published in various parts of the world which are wholly or partly devoted to the field of statistical theory and method and which will be brought within the scope of this journal of abstracts. In due course it will be possible to issue a complete list of these journals. In the case of the following journals, however, being those which are wholly devoted to statistical theory—all contributions, whether a paper, note or miscellanea, will be abstracted :

Annals of Mathematical Statistics,
Biometrika
Journal, Royal Statistical Society (Series B)
Bulletin of Mathematical Statistics
Annals, Institute of Statistical Mathematics

Within the larger group of journals, which are not wholly devoted to statistical theory and method, there are some journals which have the vast majority of their contributions in this field. These journals, therefore, will be abstracted on a virtually complete basis :

Biometrics
Metrika
Metron
Review, International Statistical Institute
Technometrics
Sankhyā

After experience in the publication of this journal of abstracts it may be found desirable to add further journals to these lists. In any case readers may be assured that all papers properly to be included in this journal of abstracts will be included irrespective of such notification. If any reader of this journal discovers a paper which happens to have been overlooked, the General Editor will be pleased to be informed so that the appropriate abstract can be made : always provided that the date of publication is after 1st October 1958, when the abstracting for this journal commenced.

In addition to the ordinary journals, there are two kinds of publication which fall within the scope of this journal of abstracts. They are the experiment and other research station reports—which occur particularly in the North American region—and the reports of conferences, symposia and seminars. Whilst these latter may be included in the book review sections of journals it is unusual for any individual contribution to be noted at any length. These publications are, in effect, special collections of papers and for this reason the appropriate arrangements will be made for them to be included in this journal. By the same token, abstracts of papers given at conferences and reproduced in an appropriate journal will be disregarded until the definitive publication is available.

FORMAT OF JOURNAL

The abstracts will be about 400 words long—the recommendation of UNESCO for the " long " abstract service : they will be in the English language although the language of the original paper will be indicated on the abstract together with the name of the abstractor. In addition, the address of the author(s) will be given in sufficient detail to facilitate contact in order to obtain further detail or request an offprint. Suitable indexes will be provided annually in a supplement.

H8327 48 Saas

The scheme of classification has been developed upon lines that will facilitate the transfer to punched cards of the code numbers allocated to each abstract: to allow for future development it is suggested that use is made of 4-column fields. Each abstract will have two classification numbers: the primary number in heavy type to indicate the basic topic of the paper and the secondary number in brackets to take account of the most important cross-reference. The sheets of the journal are colour-coded according to the twelve main sections of the classification and it should be noted that it is the main section number of the primary classification which determines the colour code for each abstract. It is believed that this method of colour-coding the pages will provide a distinctive visual aid to the identification of abstracts both when the journal is in bound form or dismantled and filed on cards or in binders. The format and simple binding allows of the following alternative treatments by users of the journal:

(*a*) Leave intact as a shelf-periodical.

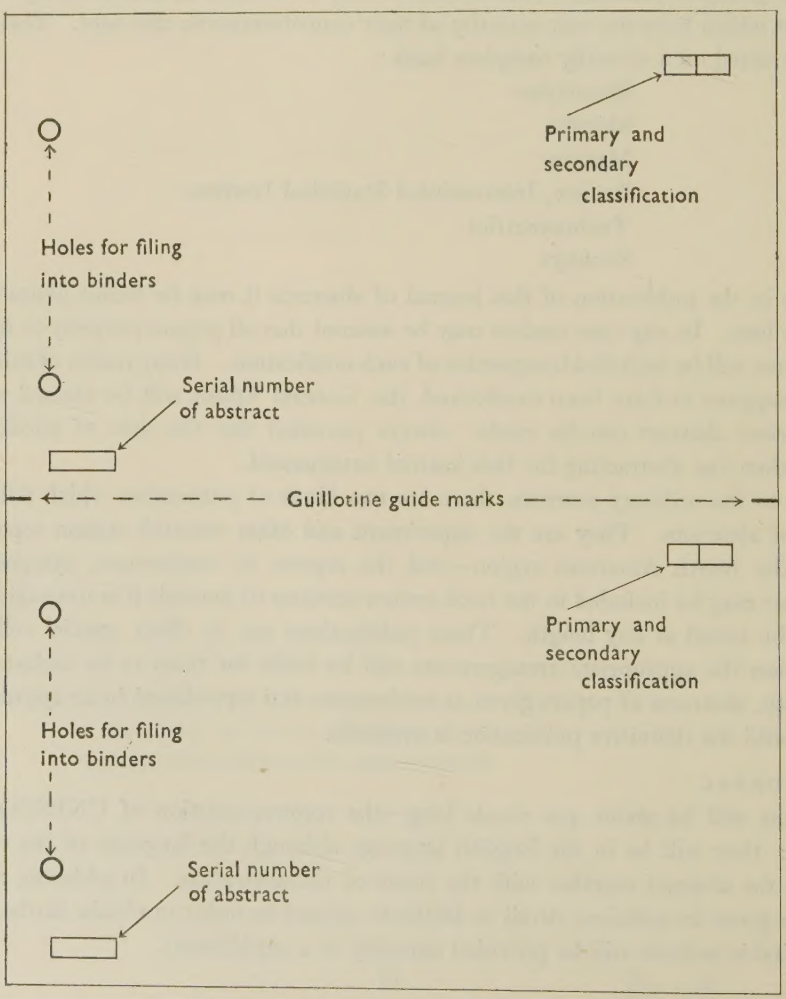
(*b*) Split and filed in page form according to the main sections of the classification.

(*c*) Split and guillotine (single cut) each page ready for:

(i) pasting onto standard index cards.

(ii) filing in loose-leaf or other binders—for which the appropriate holes are punched.

It will also be possible for those users who need a completely referenced file to insert skeleton cards or sheets according to the secondary classification number provided on each abstract.


Primary and secondary classification
Holes for filing into binders
Serial number of abstract
Guillotine guide marks
Primary and secondary classification
Holes for filing into binders
Serial number of abstract

INTERNATIONAL JOURNAL OF ABSTRACTS STATISTICAL THEORY AND METHOD

LIST OF PUBLICATIONS FOR REGULAR SCANNING

The following list of publications in the statistical and allied fields will be regularly examined for papers which should be abstracted for this Journal. Other journals will also be scrutinised from time to time and a check maintained by scanning other relevant abstracting journals. In addition, abstracts will be included of special collections of papers as published in reports of conferences, symposia and seminars together with the published reports of experiment and other research stations. The coverage of this Journal amounts to at least 400 sources of publication

Acta Acad. Cient., Lima	Acta de la Academia Cientifica, Lima	Peru
Acta Math., Stockh.	Acta Mathematica	Sweden
Acta Math. Acad. Sci. Hung.	Acta Mathematica Academiae Scientiarum Hungaricae	Hungary
Acta Math. Sinica	Acta Mathematica Sinica	China
Acta Pontif. Acad. Sci.	Acta Pontificiae Academiae Scientiarum	Italy
Advanced Management	Advanced Management	U.S.A.
Advanc. Enzymol.	Advances in Enzymology and Related Subjects	U.S.A.
Advanc. Protochem.	Advances in Protochemistry	U.S.A.
Advanc. Vet. Sci.	Advances in Veterinary Science	U.S.A.
Agric. Engng., St Joseph, Mich.	Agricultural Engineering	U.S.A.
Agriculture, Lond.	Agriculture (Min. of Agric.)	Great Britain
Agron. J.	Agronomy Journal	U.S.A.
Allg. Statist. Arch.	Allgemeines Statistisches Archiv	Germany
Amer. Econ. Rev.	American Economic Review	U.S.A.
Amer. Heart J.	American Heart Journal	U.S.A.
Amer. J. Clin. Nutr.	American Journal of Clinical Nutrition	U.S.A.
Amer. J. Clin. Path.	American Journal of Clinical Pathology	U.S.A.
Amer. J. Digestive Diseases	American Journal of Digestive Diseases	U.S.A.
Amer. J. Hum. Genet.	American Journal of Human Genetics	U.S.A.
Amer. J. Hyg.	American Journal of Hygiene	U.S.A.
Amer. J. Med.	American Journal of Medicine	U.S.A.
Amer. J. Physiol.	American Journal of Physiology	U.S.A.
Amer. J. Publ. Health	American Journal of Public Health	U.S.A.
Amer. J. Sci.	American Journal of Science	U.S.A.
Amer. J. Sociol.	American Journal of Sociology	U.S.A.
Amer. J. Vet. Res.	American Journal of Veterinary Research	U.S.A.
Amer. Nat.	American Naturalist	U.S.A.
Amer. Statistician	American Statistician	U.S.A.
Analyt. Chem.	Analytical Chemistry	U.S.A.
An. Soc. Cient. Argent.	Anales de la Sociedad Cientifica Argentina	Argentina
Ann. Acad. Sci. Fenn.	Annales Academiae Scientiarum Fennicae	Finland
Ann. Appl. Biol.	Annals of Applied Biology	Great Britain
Ann. Fac. Econ. Com. Catania	Annali della Facoltà di Economia e Commercio, Catania (University)	Italy
Ann. Fac. Econ. Com. Palermo	Annali della Facoltà di Economia e Commercio, Palermo (University)	Italy
Ann. Fac. Sci. Porto	Annaes da Faculdade de Ciências do Porto	Portugal
Ann. Hum. Genet.	Annals of Human Genetics (London)	Great Britain
Ann. Inst. Poincaré	Annales de l'Institut Henri Poincaré	France
Ann. Inst. Statist. Bari	Annali dell'Instituto di Statistica, Bari (University)	Italy
Ann. Inst. Statist. Math., Tokyo	Annals of the Institute of Statistical Mathematics	Japan
Ann. Inst. Sup. Sci. Econ. Financ.	Annaes do Instituto Superior de Ciências Economicas e Financieras	Portugal
Ann. Math. Statist.	Annals of Mathematical Statistics	U.S.A.
Ann. N.Y. Acad. Sci.	Annals of the New York Academy of Science	U.S.A.
Ann. Sci. Éc. Norm. Sup., Paris	Annales Sciéntifiques de l'École Normale Supérieure	France
Ann. Sci. Écon. Appl.	Annales de Sciences Économique Appliquées	France
Ann. Soc. Sci. Bruxelles, I	Annales de la Société Scientifique de Bruxelles—Serie 1	Belgium

Ann. Statist., Rome	Annali de Statistica (Instituto Centrale di Statistica, Rome)	Italy
Ann. Univ. Lyon, A	Annales de l'Université de Lyon—Section A	France
Ann. Univ. Sci. Budapest. (Sect. Math.)	Annales Universitatis Scientiarum Budapestiensis de Rolando Eötvös nominatae—Sectio Mathematica	Hungary
Annu. Rev. Physiol.	Annual Review of Physiology	U.S.A.
Aplik. Mat.	Aplikace Matematiky	Czechoslovakia
Appl. Sci. Res., Hague	Applied Scientific Research, The Hague	Netherlands
Appl. Statist.	Applied Statistics	Great Britain
Archimede	Archimede	Italy
Arch. Biochem. Biophys.	Archives of Biochemistry and Biophysics	U.S.A.
Arch. Math., Karlsruhe	Archiv der Mathematik	Germany
Ark. Mat.	Arkiv för Matematik	Sweden
Atti (mem.) Accad. Lincei	Atti della Accademia Nazionale dei Lincei	Italy
Atti Riun. Sci. Soc. Ital. Statist.	Atti della Riunione Scientifica Società Italiana di Statistica, Rome	Italy
Aust. J. Agric. Res.	Australian Journal of Agricultural Research	Australia
Aust. J. Appl. Sci.	Australian Journal of Applied Science	Australia
Aust. J. Biol. Sci.	Australian Journal of Biological Sciences	Australia
Aust. J. Phys.	Australian Journal of Physics	Australia
Aust. J. Sci.	Australian Journal of Science	Australia
Aust. J. Statist.	Australian Journal of Statistics	Australia
Beil. Mber. Öst. Inst. Wirtschaftsf.	Beilagen zu den Monatsberichten des Österreichischen Institutes für Wirtschaftsforschung	Austria
Bell Lab. Record	Bell Laboratories Record	U.S.A.
Bell Syst. Monographs	Bell System Monographs	U.S.A.
Bell Syst. Tech. J.	Bell System Technical Journal	U.S.A.
Biometrics	Biometrics	U.S.A.
Biometrika	Biometrika	Great Britain
Biom. Zeit.	Biometrische Zeitschrift	Germany
Blä. Dtsch. Ges. Versich.-math.	Blätter der Deutschen Gesellschaft für Versicherungsmathematik	Germany
Bol. Fac. Filos. Ciênt., S. Paulo	Boletim da Faculdade de Filosofía, Ciências e Letras, São Paulo	Brazil
Bol. Inst. Actuar. Port.	Boletim do Instituto dos Actuarios Portugueses	Portugal
Bol. Soc. Mat. Mexicana	Boletim de la Sociedad Matemática Mexicana	Mexico
Boll. Cent. Ric. Operat.	Bollettino del Centro per la Ricerca Operativa	Italy
Boll. Un. Mat. Ital.	Bollettino dell'Unione Matematica Italiana	Italy
Brit. J. Nutr.	British Journal of Nutrition	Great Britain
Brit. J. Prev. Soc. Med.	British Journal of Preventive and Social Medicine	Great Britain
Brit. J. Sociol.	British Journal of Sociology	Great Britain
Brit. J. Statist. Psychol.	British Journal of Statistical Psychology	Great Britain
Bull. Acad. Sci., III	Bulletin de l'Académie des Sciences (Classe 3)	Poland
Bull. Amer. Soc. Test. Mat.	Bulletin, American Society for Testing Materials	U.S.A.
Bull. Ass. Actuair. Suisses	Bulletin de l'Association des Actuaires Suisses	Switzerland
Bull. Calcutta Math. Soc.	Bulletin of the Calcutta Mathematical Society	India
Bull. Calcutta Statist. Ass.	Bulletin, Calcutta Statistical Association	India
Bull., C.S.I.R.O., Aust.	Bulletins, Commonwealth Scientific and Industrial Research Organisation	Australia
Bull. Indian Soc. Qual. Contr.	Bulletin, Indian Society of Quality Control	India
Bull. Indian Stand. Inst.	Bulletin, Indian Standards Institution	India
Bull. Inst. Égypte	Bulletin de l'Institut d'Égypte, Cairo	Egypt
Bull. Jap. Statist. Soc.	Bulleting of the Japan Statistical Society	Japan
Bull. Math. Statist.	Bulletin of Mathematical Statistics, Research Association of Statistical Sciences	Japan
Bull. Nat. Inst. Sci. India	Bulletin, National Institute of Sciences of India	India
Bull. Oxf. Univ. Inst. Statist.	Bulletin of the Oxford University Institute of Statistics	Great Britain
Bull. Res. Inst. Kerala	Bulletin, Central Research Institute, Kerala (University)	India
Bull. Sci. Math.	Bulletin des Sciences Mathématique	France
Bull. Soc. Math. Belg.	Bulletin de la Société Mathématique de Belgique, Gembloux	Belgium

Bull. Soc. Math. France	Bulletin de la Société Mathématique de France	France
Bull. Statist. Soc., N.S.W.	Bulletin of the Statistical Society, N.S. Wales	Australia
Canad. J. Econ. Polit. Sci.	Canadian Journal of Economic and Political Science	Canada
Cancer, New York	Cancer (New York)	U.S.A.
Cas. pěst. math.	Casopis pro pěstování mathematiky	Czechoslovakia
Chem. Engng. Progr.	Chemical Engineering Progress	U.S.A.
Clin. Chem.	Clinical Chemistry	U.S.A.
Colect. Estud.	Colectanea de Estudos	Spain
Coll. Mat.	Collectanea Matematica	Spain
Colloq. Math.	Colloquium Mathematicum	Poland
Comment. Math. Helvetia	Commentarii Mathematica Helvetia	Switzerland
Comment. Pontif. Acad. Sci.	Commentationes Pontificiae Academiae Scientiarum	Italy
Compos. Math., Groningen	Compositio Mathematica, Groningen	Netherlands
C.R. Acad. Sci., Paris	Comptes Rendus de l'Académie des Sciences, Paris	France
Cornell Vet.	Cornell Veterinarian	U.S.A.
Current Science	Current Science	India
Curso Mira Fernandes	Curso Mira Fernandes	Spain
Czech. Math. J.	Czechoslowak Mathematical Journal	Czechoslovakia
Dacca Univ. Statist. Bull.	Dacca University Statistical Bulletin	Pakistan
Defense Sci. J.	Defense Science Journal	India
Dokl. Akad. Nauk. SSSR	Doklady Akademii Nauk, SSSR.	U.S.S.R.
Dopov. Akad. Nauk. Ukrain, SSR	Dopovedy Akademii Nauk Ukrainskoj, SSR	U.S.S.R.
Écon. Appl.	Économie Appliquée	France
Econ. Internatzionale	Economia Internatzionale	Italy
Econ. J.	Economic Journal	Great Britain
Écon. Leban. Arabe	L'Économie Lebanaise & Arabe, Beirut	Lebanon
Econ. Record	Economic Record	Australia
Econometrica	Econometrica	U.S.A.
Economica	Economica	Great Britain
Edinb. Math. Notes	Edinburgh Mathematical Notes	Great Britain
Égypte Contemp.	L'Égypte Contemporaine, Cairo	Egypt
Egyptian Statist. J.	Egyptian Statistical Journal	Egypt
Ekon. Tidskr.	Ekonomisk Tidskrifts	Sweden
Emp. J. Exp. Agric.	Empire Journal of Experimental Agriculture	Great Britain
Endocrinology	Endocrinology	U.S.A.
Engineer, Lond.	Engineer	Great Britain
Engineering, Lond.	Engineering	Great Britain
Ergebn. Naturw. Grenzgeb.	Ergebnisse der Naturwissenschaften und ihrer Grenzgebiete	Germany
Esta	Esta	Spain
Estadistica	Estadistica	U.S.A.
Euclides, Groningen	Euclides, Groningen	Netherlands
Eugen. Rev.	Eugenics Review	Great Britain
Evolution	Evolution	U.S.A.
Food Res.	Food Research	U.S.A.
Food Tech.	Food Technology	U.S.A.
Ganita	Ganita, Lucknow-Allahabad (Universities)	India
Gaz. Mat.	Gazeta de Matematica	Spain
G. Economisti	Giornale degli Economisti e Annali di Economia	Italy
G. Ist. Ital. Attuari	Giornale dell'Istituto Italiano degli Attuari	Italy
Genetics	Genetics	U.S.A.
Genus	Genus	Italy
Growth	Growth	U.S.A.
Harvard Busin. Rev.	Harvard Business Review	U.S.A.
Helvetica Phys. Acta	Helvetica Physica Acta	Switzerland

Heredity	Heredity	Great Britain
Hum. Biol.	Human Biology	U.S.A.
IFO-Studien	IFO-Studien	Austria
Inc. Statist.	Incorporated Statistician	Great Britain
Indag. Math.	Indagationes Mathematicae, Amsterdam	Netherlands
Indian Coun. Agric. Res. Bull.	Indian Council of Agricultural Research : Bulletin	India
Indian Coun. Agric. Res. Statist. Newsletter	Indian Council of Agricultural Research : Statistical Newsletter	India
Indian Econ. J.	Indian Economic Journal	India
Indian Econ. Rev.	Indian Economic Review	India
Indian J. Agric. Econ.	Indian Journal of Agricultural Economics	India
Indian J. Phys.	Indian Journal of Physics	India
Indian J. Power & River Devel.	Indian Journal of Power and River Development	India
Indian J. Soc. Work	Indian Journal of Social Work	India
Indian Math. Soc. J.	Indian Mathematical Society Journal	India
Industr. Engng. Chem.	Industrial and Engineering Chemistry	U.S.A.
Industr. Qual. Contr.	Industrial Quality Control	U.S.A.
Industria, Milano	L'Industria	Italy
Iowa Acad. Sci. J.	Iowa Academy of Science Journal	U.S.A.
Iowa Exp. Sta. Publ.	Iowa State College Experiment Station Publications	U.S.A.
Izv. Akad. Nauk, SSSR	Izvestiya Akademii Nauk, SSSR	U.S.S.R.
Izv. Akad. Nauk, Uzbek., SSR.	Izvestiya Akademii Nauk, Uzbekskoj, SSR.	U.S.S.R.
Jb. Nat.-Ökon. Statist.	Jahrbucher für Nationalökonomie und Statistik	Germany
Jber. Dtsch. Mathver.	Jahresbericht der Deutschen Mathematikervereinigung	Germany
J. Agric. Food Chem.	Journal of Agricultural and Food Chemistry	U.S.A.
J. Agric. Sci.	Journal of Agricultural Science	Great Britain
J. Amer. Pharm. Ass.	Journal of the American Pharmaceutical Association (Scientific Edition)	U.S.A.
J. Amer. Statist. Ass.	Journal, American Statistical Association	U.S.A.
J. Amer. Vet. Med. Ass.	Journal of American Veterinary Medicine Association	U.S.A.
J. Animal Sci.	Journal of Animal Sciences	U.S.A.
J. Appl. Physiol.	Journal of Applied Physiology	U.S.A.
J. Ass. Comp. Mach.	Journal of the Association of Computing Machinery	U.S.A.
J. Aust. Inst. Agric. Sci.	Journal of the Australian Institute of Agricultural Science	Australia
J. Aust. Math. Soc.	Journal of the Australian Mathematical Society	Australia
J. Biol. Chem.	Journal of Biological Chemistry	U.S.A.
J. Chronic Diseases	Journal of Chronic Diseases	U.S.A.
J. Clin. Invest.	Journal of Clinical Investigation	U.S.A.
J. Dairy Sci.	Journal of Dairy Science	U.S.A.
J. Ecology	Journal of Ecology	Great Britain
J. Econ. Ent.	Journal of Economic Entomology	U.S.A.
J. Exp. Educ.	Journal of Experimental Education	U.S.A.
J. Exp. Med.	Journal of Experimental Medicine	U.S.A.
J. Forestry	Journal of Forestry	U.S.A.
J. Gen. Physiol.	Journal of General Physiology	U.S.A.
J. Genet.	Journal of Genetics	India
J. Heredity	Journal of Heredity	U.S.A.
J. Hyg., Camb.	Journal of Hygiene, Cambridge	Great Britain
J. Industr. Econ.	Journal of Industrial Economics	Great Britain
J. Industr. Engng.	Journal of Industrial Engineering	U.S.A.
J. Indian Soc. Agric. Statist.	Journal of the Indian Society of Agricultural Statistics	India
J. Inst. Actuar.	Journal of the Institute of Actuaries	Great Britain
J. Inst. Actuar. Stud. Soc.	Journal of the Institute of Actuaries Students' Society	Great Britain
J. Madras Uuiv., B	Journal of Madras University, Section B.	Great Britain
J. Mammalogy	Journal of Mammalogy	U.S.A.
J. Marine Res.	Journal of Marine Research	U.S.A.
J. Math. Pures Appl.	Journal de Mathématiques Pures et Appliquées	France
J. Math. Soc. Japan	Journal of the Mathematical Society of Japan	Japan
J. Nutrition	Journal of Nutrition	U.S.A.
J. Operat. Res.	Journal of the Operations Research Society of America	U.S.A.

J. Pharmacol.	Journal of Pharmacology and Experimental Therapeutics	U.S.A.
J. Plant Physiol.	Journal of Plant Physiology	U.S.A.
J. Polit. Econ.	Journal of Political Economy	U.S.A.
J. reine angew. Math.	Journal für reine und angewandte Mathematik	Germany
J. Res., Nat. Bur. Stand.	Journal of Research, National Bureau of Standards, Washington	U.S.A.
J. Roy. Soc., N.S.W.	Journal of the Royal Society, N.S. Wales	Australia
J. R. Statist. Soc. (A & B)	Journal of the Royal Statistical Society (Series A and B)	Great Britain
J. Sci. Engng.	Journal of Science and Engineering	India
J. Sci. Industr. Res.	Journal of Scientific and Industrial Research	U.S.A.
J. Sci. Industr. Res., India	Journal of Scientific and Industrial Research	India
J. Soc. Statist. Paris	Journal de la Société Statistique de Paris	France
J. Statist. Social Inquiry Soc.	Journal of the Statistical and Social Inquiry Society of Ireland	Ireland
J. Univ. Baroda	Journal, Maharaja Sayajiroa, Baroda University	India
Kodai (Tokyo Inst. Tech.)	Kodai (Mathematical Seminar reports) Tokyo Institute of Technology	Japan
Magy. Tud. Akad. III *Oszl. Kozl.*	Magyar Tudományos Akadémia III Matematika és Fizikai Oszlályának Kozleményei	Hungary
Management Sci.	Management Science	U.S.A.
Manager (India)	Manager	India
Manchester School	Manchester School	Great Britain
Matemática	Matemática, Notas del Instituto de Matemática Facultad de Ingeniería de Rosario	Argentine
Mat. Lapok.	Matematikai Lapok	Hungary
Math. Ann.	Mathematische Annalen	Germany
Math. Gaz.	Mathematical Gazette	Great Britain
Math. Nachr.	Mathematische Nachrichten	Germany
Math. Stud.	Mathematics Students	India
Math. Tab., Wash.	Mathematical Tables and other aids to Computation	U.S.A.
Math., Tech., Wirtschaft	Mathematik, Technik, Wirtschaft (Zeitschrift für moderna Rechentechnik und Automation)	Germany
Math. Zeit.	Mathematische Zeitschrift	Germany
Meded. Landbhoogesch., Wageningen	Mededeelingen van de Landbouwhoogeschool, Wageningen	Netherlands
Mem. Coll. Sci., Kyoto	Memoirs of the College of Science, Kyoto (University)	Japan
Mem. Fac. Sci., A, *Kyusyu*	Memoirs of the Faculty of Science (Series A), Kyusyu (University)	Japan
Mém. Soc. Sci. Liége	Mémoires de la Société Royale des Sciences de Liége	Belgium
Metabolism	Metabolism	U.S.A.
Met. Mag., Lond.	Meteorological Magazine	Great Britain
Metrika	Metrika	Germany
Metroeconomica	Metroeconomica	Italy
Metron	Metron	Italy
Mitt. Ver. Schweiz. Versich.-Math.	Mitteilungen der Vereinigung Schweizerisher Versicherungs-Mathematiker	Switzerland
Monatsh. Math.	Monatsheft für Mathematik	Germany
Monthly Not. R. Astr. Soc.	Monthly Notices of the Royal Astronomical Society	Great Britain
Nat. Bur. Stand., Statist. Rep. & Publ.	National Bureau of Standards, Washington : Statistical Reports and Publications	U.S.A.
Nature	Nature (London)	Great Britain
Ned. Verzek. Actuar. Bij.	Nederlands Verzekeringsarchief, Actuarieel Bijvoegsel, The Hague	Netherlands
Nieuw Arch. Wisk.	Nieuw Archief voor Wiskunde, Groningen	Netherlands
Nieuw Tijdschr. Wisk.	Nieuw Tijdschrift voor Wiskunde, Groningen	Netherlands

Operat. Res. Quart.	Operational Research Quarterly	Great Britain
Osaka Math. J.	Osaka Mathematical Journal, Osaka (University)	Japan
Öst. Ingen.-arch.	Österreichisches Ingenieurarchiv	Austria
Oxf. Econ. Papers	Oxford Economic Papers	Great Britain
Pakistan Econ. J.	Pakistan Economic Journal	Pakistan
Phil. Mag.	Philosophical Magazine	Great Britain
Phil. Phenom. Res.	Philosophical and Phenomenological Research	Great Britain
Phil. Trans.	Philosophical Transactions of the Royal Society	Great Britain
Population Stud.	Population Studies	Great Britain
Portug. Math.	Portugaliae Mathematica	Portugal
Poultry Sci.	Poultry Science	U.S.A.
Proc. Amer. Soc. Hort. Sci.	Proceedings, American Society for Horticultural Science	U.S.A.
Proc. Amer. Soc. Qual. Contr.	Proceedings of the American Society for Quality Control	U.S.A.
Proc. Amer. Soc. Test. Mat.	Proceedings of the American Society for Testing Materials	U.S.A.
Proc. Camb. Phil. Soc.	Proceedings of the Cambridge Philosophical Society	Great Britain
Proc. Indian Acad. Sci.	Proceedings of the Indian Academy of Sciences	India
Proc. Indian Ass. Cult. Sci.	Proceedings of the Indian Association for Cultivation of Science	India
Proc. Indian Sci. Congr.	Proceedings of the Indian Science Congress	India
Proc. Inst. Statist. Math.	Proceedings of the Institute of Statistical Mathematics	Japan
Proc. Japan Acad.	Proceedings of the Japan Academy	Japan
Proc. Kon. Ned. Akad., Wetensch. A	Proceedings, Koninklijke Nederlandsche Akademie van Wetenschappen (A)	Netherlands
Proc. Lahore Phil. Soc.	Proceedings of the Lahore Philosophical Society	India
Proc. Lond. Math. Soc.	Proceedings of the London Mathematical Society	Great Britain
Proc. Math. Phys. Soc. Egypt	Proceedings of the Mathematical and Physical Society of Egypt	Egypt
Proc. Nat. Acad. Sci. India	Proceedings of the National Academy of Sciences	India
Proc. Nat. Inst. Sci. India (A & B)	Proceedings of the National Institute of Sciences of India (Section A and B)	India
Proc. Pakistan Statist. Ass.	Proceedings of the Pakistan Statistical Association	Pakistan
Proc. Qual. Contr. Ass., Bangalore	Proceedings of the Quality Control Association, Bangalore	India
Proc. R. Irish Acad.	Proceedings of the Royal Irish Academy	Ireland
Proc. Roy. Soc. A & B	Proceedings of the Royal Society (Series A & B)	Great Britain
Proc. Roy. Soc. N.S.W.	Proceedings of the Royal Society of N.S. Wales	Australia
Proc. Roy. Soc. N.Z.	Proceedings of the Royal Society of New Zealand	New Zealand
Proc. Soc. Exp. Biol., N.Y.	Proceedings of the Society of Experimental Biology and Medicine	U.S.A.
Proc. Soil Sci. Soc. Amer.	Proceedings, Soil Science Society of America	U.S.A.
Przegl. Statyst.	Przeglad Statystyczny	Poland
Psychometrika	Psychometrika	U.S.A.
Public Health Rep.	Public Health Reports	U.S.A.
Public Opinion Quart.	Public Opinion Quarterly	U.S.A.
Publ. Inst. Mat. Estadíst. Fac. Ingen. Agric., Montevideo	Publicaciones del Instituto de Matemáticas y Estadística de la Facultad de Ingenieria y Agrimensura, Montevideo	Uruguay
Publ. Inst. Statist. Paris	Publications de l'Institut de Statistique de l'Université de Paris	France
Publ. Math., Debrecen	Publicationes Mathematicae	Hungary
Publ. Math. Inst. Hung. Acad. Sci.	Publications of the Mathematical Institute of the Hungarian Academy of Sciences	Hungary
Quart. J. Econ.	Quarterly Journal of Economics	U.S.A.
Quart. J. Exp. Physiol.	Quarterly Journal of Experimental Physiology	U.S.A.
Quart. J. Math.	Quarterly Journal of mathematics, Oxford	Great Britain
Quart. J. Med.	Quarterly Journal of Medicine	Great Britain
Quart. J. R. Met. Soc.	Quarterly Journal of the Royal Meteorological Society	Great Britain

RAND Corp., Publ. & Rep.	RAND Corporation, Publications and Reports	U.S.A.
R.C. Accad. Lincei	Rendiconti delle Accademia Nazionale dei Lincei	Italy
R.C. Circ. Mat.	Rendiconti del Circolo Matematico	Italy
R.C. Mat. Univ. Roma.	Rendiconti di Matematica e delle sue applicazioni	Italy
R.C. Semin. Sci. Cagliari	Rendiconti del Seminario della Facoltà di Scienze, Cagliari (University)	Italy
Rep. Statist. Appl. Res. (JUSE)	Reports of Statistical Application Research, Union of Japanese Scientists & Engineers	Japan
Research, Beirut	Research, A.U.B.—Beirut	Lebanon
Rev. Acad. Madrid	Revista de la Real Academia de Ciencias Exactas	Spain
Rev. Brasil. Estatist.	Revista Brasileira de Estatística	Brazil
Rev. Cienc. Tucuman	Revista de Ciencias de la Universidad Nacional del Tucumán	Argentine
Rev. Écon.	Revue Économique	France
Rev. Econ. Statist.	Review of Economics and Statistics	U.S.A.
Rev. Econ. Stud.	Review of Economic Studies	Great Britain
Rev. Economia, Lisbon	Revista de Economia	Portugal
Rev. Estocást. Soc. Argent. Estadíst.	Revista Estocástica de la Sociedad Argentina de Estadística	Argentine
Rev. Fac. Cienc. Econ. Admin. Montevideo	Revista de la Facultad de Ciencias Economicas y de Administracion, Montevideo	Uruguay
Rev. Int. Statist. Inst.	Review, International Statistical Institute	Netherlands
Rev. Ital. Econ. Demogr. Statist.	Revista Italiana di Economia Demografia e Statistica	Italy
Rev. Mat. Hisp.-Amer.	Revista Matematica Hispano-Americana	Spain
Rev. Rech. Operat.	Revue de Recherche Operationelle	France
Rev. Sci., Paris	Revue Scientifique	France
Rev. Soc. Cubana Cienc. Fìs. Mat.	Revista de la Sociedad Cubana de Ciencas, Fìsicas y Matemàticas	Cuba
Rev. Statist. Appl.	Revue de Statistique Appliquée	France
Ric. Sci.	Ricerca Scientifica	Italy
Riv. Int. Sci. Soc.	Rivista Internazionale di Science Sociali	Italy
Riv. Polit. Econ.	Rivista di Politica Economica	Italy
Sankhyā	Sankhyā ; Indian Journal of Statistics	India
Schmollers Jb.	Schmollers Jahrbuch	Germany
Schr. Ver. Sozialpolit.	Schriften des Vereins für Sozialpolitik	Germany
Schweiz. Zeit. Volkwirtsch. Statist.	Schweizerische Zeitschrift für Volkwirtschaft und Statistik	Switzerland
Science	Science	U.S.A.
Science & Culture	Science and Culture	India
Scientia	Scientia	Italy
Scottish J. Polit. Econ.	Scottish Journal of Political Economy	Great Britain
Simon Stevin	Simon Stevin, Groningen	Netherlands
Skand. Aktuartidskr.	Skandinavisk Aktuarietidskrift	Sweden
Sociometry	Sociometry	U.S.A.
Soil Sci.	Soil Science	U.S.A.
South Africa J. Econ.	South African Journal of Economics	South Africa
Statistica, Bologna	Statistica	Italy
Statist. Nachr.	Statistische Nachrichten	Germany
Statist. Neerlandica	Statistica Neerlandica, The Hague	Netherlands
Statist. Praxis	Statistische Praxis	Germany
Statist. Qual. Contr.	Statistical Quality Control	Japan
Statist. Rev. Belgrade	Statisticka Revija	Yugoslavia
Studi. Mercato	Studi di Mercato	Italy
Studia Math.	Studia Mathematica	Poland
Sugaku	Sugaku (Mathematical Society of Japan)	Japan
Tech. News, Nat. Bur. Stand.	Technical News Letter, National Bureau of Standards, Washington	U.S.A.
Tech. Publ., Amer. Soc. Test. Mat.	Technical Publications, American Society for Testing Materials	U.S.A.
Tech. Rep., C.S.I.R.O., Aust.	Technical Reports, Commonwealth Scientific & Industrial Research Organisation	Australia

Technometrics	Technometrics	U.S.A.
Teor. Veroyat. Primen.	Teorija Veroyatnosteyi Primenenia	U.S.S.R.
Textile Res.	Textile Research	U.S.A.
Tobacco Sci.	Tobacco Science	U.S.A.
Trab. Estadíst.	Trabajos de Estadística	Spain
Trans. Edin. Math. Soc.	Transactions of the Edinburgh Mathematical Society	Great Britain
Trans. Fac. Actuar., Edinb.	Transactions of the Faculty of Actuaries	Great Britain
Trans. Manchester Statist. Soc.	Transactions of the Manchester Statistical Society	Great Britain
Trans. Nat. Inst. Sci. India	Transactions of the National Institute of Sciences of India	India
Trans. N.Y. Acad. Sci.	Transactions of the New York Academy of Science	U.S.A.
Trans. Roy. Soc. N.Z.	Transactions of the Royal Society, N. Zealand	New Zealand
Trans. Soc. Actuar., Chicago	Transactions of the Society of Actuaries	U.S.A.
U krain. Mat. Z.	Ukrainskij Matematiceskij Zurnal	U.S.S.R.
U.N. (F.A.O.) Tech. Publ.	United Nations (Food and Agricultural Organisation) Technical Publications	Italy
U.N. Statist. Publ.	United Nations : Statistical Office Publications	U.S.A.
Unternehmensforschung (O.R.)	Unternehmensforschung (Operations Research)	Germany
U.S. Census Publ.	U.S. Bureau of the Census, Publications	U.S.A.
U.S. Dept. Agric., Publ.	U.S. Department of Agriculture, Publications	U.S.A.
Verh. Kon. Belg. Akad., II	Verhandelingen. Koninklijke Belgische Akademie (Serie 2)	Belgium
Verh. Vlaamse Akad. Belg.	Verhandelingen. Koninklijke Vlaamsche Akademie	Belgium
Vitamins & Hormones	Vitamins and Hormones	U.S.A.
Weltwirts. Arch.	Weltwirtschaftliches Archiv	Germany
Wirtsch. und Statist.	Wirtschaft und Statistik	Germany
Yorkshire Bull. Econ. Soc. Res.	Yorkshire Bulletin of Economic and Social Research	Great Britain
Zastosowania Mat.	Zastosowania Matematyki	Poland
Zeit. angew. Math. Mech.	Zeitschrift für angewandete Mathematik und Mechanik	Germany
Zeit. ges. Staatswiss.	Zeitschrift für die gesamte Staatswissenschaft	Germany
Zeit. Handelswiss. Forsch.	Zeitschrift für Handelswissenschaftliche Forschung	Germany
Zeit. Nat.-Ökon.	Zeitschrift für Nationalökonomie	Germany

The short titles given above are based upon the principles stated in the League of Nations report *International Code of Abbreviations for Titles of Periodicals* (1930) and its *Supplement* (1932) as also used in the *World List of Scientific Periodicals*

ABRHAM J. (Charles University, Prague) **0.8** (0.1)

An approximate method for non-linear programming—*In Czech*
Čas. pěst. math. (1958) **83**, 425-439 (5 refs.)

The paper gives a solution of the problem of finding the minimum of a strictly convex function on the set of all non-negative solutions of a system of linear equations. This is a generalisation of the basic problem of linear programming.

The author constructs an iterative process for solving the problem and shows that it converges to the true solution. He shows further how this method can be applied to quadratic programming. He discusses a special case where the function to be minimised is a sum of a linear form and a positive definite quadratic form and shows that in this case the iterative process can often be modified so that it stops after a finite number of steps. In the last section of the paper the author gives a numerical example with some practical advice for getting the solution more quickly.

The iterative process formulated by the author is the following : let us have a system of $m+k$ linear equations in m unknown quantities, let M be the set of all non-negative solutions of this system. By a vector of the set M shall be meant a solution of the corresponding homogeneous system. For every point X in M there can be chosen a vector basis $(u_1, \ldots, u_k)$ satisfying the following conditions : (*a*) the vectors u_j are linearly independent, (*b*) any u_j has zeros at all or at all but one components which are zero for X, (*c*) any component which is zero for X is non-zero for exactly one u_j.

Let f be the strictly convex function we have to minimise on M. We start with an arbitrary point X_1 of M and choose a corresponding basis $(u_1, \ldots, u_k)$ and then a point $X_2 = X_1 + tu_1$ such that $f(X_2)$ be minimum under the condition that $X_2 \in M$. If the given basis is a basis for X_2, too, we seek $X_3 = X_2 + tu_2$ in the same manner, and we continue so for $X_4, \ldots$, taking the vectors u_j cyclically one after the other. If the given basis cannot be used for the new point, a new basis is chosen and we proceed as when starting with X_1.

(F. Zítek)

AKAIKE, H. (Institute of Statistical Mathematics, Tokyo) **0.6** (6.4)

On a computation method for eigenvalue problems and its application to statistical analysis—*In English*
Ann. Inst. Statist. Math., Tokyo (1958) **10**, 1-20 (8 refs., 2 tables)

The paper deals with the numerical method for solving the N-dimensional eigenvalue problem $Ax = \lambda Bx$: where A is a real symmetric matrix and B a positive definite real symmetric matrix, both N by N, and x is an N-dimensional column vector (N being large). The paper commences with a discussion of a general approximation procedure followed by individual cases : first for one vector ($k = 1$) and then for two specified forms of the functions ξ. A computation scheme is presented which is cyclic and converging with checks for the errors of rounding. In the actual computation a finite digital representation of each quantity has to be made and because of this the problem sometimes reduces to a linear form after some steps in the approximation. The paper proceeds to discuss an accelerated scheme with particular regard to the convergence property and rounding errors. The theoretical exposition of the papers concludes with a generalisation of the scheme—after noting the convergence to the eigenvector with the maximum eigenvalue, and that in practical applications it is frequently necessary only to determine this single eigenvector.

The scheme of computation given in this paper for a high-speed computer and the numerical examples given in the paper were all worked out on an automatic relay computer and the technical details will be the subject of a further paper. The paper concludes with some statistical applications in the field of multivariate analysis : canonical correlations, the linear discriminant function, the quantification problem and trend analysis of multiple time series. For canonical correlation the ordinary formulation as well as the least squares method is briefly considered.

(C. Hayashi)

COX, C. P. (Nat. Inst. for Res. in Dairying, Univ. of Reading)
A concise derivation of general orthogonal polynomials—*In English*
J. R. Statist. Soc. B (1958) **20**, 406-407

0.6 (6.1)

Orthogonal polynomials for describing the regression of y and x are easily derived from first principles. The author gives an alternative derivation by generalising the polynomials from the general regression equation in its determinantal form using the Schweinsian expansion of a determinant. The partitioning of a sum of squares into independent parts corresponding to successively fitted orthogonal polynomials is also derived in a similar way.

(F. N. David)

FÜRST, D. (University of Rome)
An application of the Theory of Games—*In Italian*
Archimede (1958) **10**, 211-223 (4 refs., 4 figs.)

0.7 (0.8)

The author gives the principles and the basic methods of the theory of games in an elementary form. Starting from a particular example (a three by three game), he introduces the general problem of two-person zero-sum games and, after showing its links with linear programming, describes the method of solving it. The third section is devoted to the solution and discussion of an example. After solving the game, the author studies how the value of the game and the optimum strategy vary when the pay-off matrix changes according to given rules.

The problem is also treated from a geometrical point of view, and this gives to the author the opportunity of pointing out the geometrical interpretation of the fundamental theorem of Von Neumann.

The last part of the paper is devoted to studying the case in which a player can modify, within certain restraints, the pay-off matrix. Since the example examined in the paper is of this type, the author treats the problem from the point of view of that player, i.e. finding the pay-off matrix for which the value of the game is a maximum.

(G. Dall'Aglio)

GOOD, I. J. (Royal Naval Scientific Service, Cheltenham) 0.5 (9.2)

The interaction algorithm and practical fourier analysis—*In English*
J. R. Statist. Soc. B (1958) **20**, 361-372 (9 refs.)

The interactions of a 2^n factorial experiment can be calculated using a simple algorithm suggested by Yates. Box suggested an extension of the algorithm of Yates for the 3^n experiment. The author points out that these algorithms imply the existence of linear formulae for the interactions and proves a theorem whereby these formulae can be written down.

He considers the situation where the interactions are known and it is desired to " recover " the observations. An inverse algorithm is given for this reverse process. The uses of this inverse algorithm are discussed. Apart from the obvious use that it enables the forward calculation of the interactions to be checked the author suggests that non-significant high order interactions may be put equal to zero and the inverse algorithm used to smooth the observations with the effect of making them closer to those expected on hypothesis.

(F. N. David)

KABE, D. G. (Karnatak University, Dharwar, India) 0.4 (3.9)

Applications of Meijer-G functions to distribution problems in statistics—*In English*
Biometrika (1958) **45**, 578-580 (5 refs.)

The author points out that, by contrast with the extensive use of the Fourier transformation in statistical distribution theory, the Mellin transformation appears to have been neglected. He notes that the Mellin transformation appears particularly useful in those cases where the moments of a statistic can be expressed in terms of the product and the ratio of Γ-functions. As illustration he derives the distribution of the generalised correlation ratio and the joint distribution of two variables, each a function of sample canonical coefficients, given in 1939 by Girskick [*Ann. Math. Statist.*, **10**, 203].

(F. N. David)

LEVI, E. (Inst. of Financial Mathematics, University of Catania)
Fitting of life tables by sums of exponential functions—*In Italian*
Statistica (1959) **19**, 3-19 (15 refs.)

0.3 (0.2)

The problem involving the theory of life table functions in connection with values of life annuities by means of the values of annuities-certain has recently been treated by several authors. The present method establishes a necessary and sufficient condition under which the l_x function is expressed as the general solution of an homogeneous linear differential equation of the kth order with constant coefficients. The above solution is given by

$$l_x = \sum_{r=0}^{k-1} b_r u_r^x$$

where the b_r are constants and the u_r are real or complex : the cases for which the u_r are all real are known.

Some mathematicians—cited by the author—are interested in suggesting particular applications of the expression given above which show some convenient methods for determining the parameters. To overcome difficulties concerning the method of least squares, the author follows two procedures : one based on " points " and another based on " sums ", that is to say " interpolation by points " and " interpolation by sums ".

By analogy with the Steffensen system of calculating life annuities, it is possible to unite the two systems in one. In the above unified system of equations the determinant of coefficients is the well-known Hankel determinant. In order to assist the solution, the author suggests a useful transformation of this determinant. Furthermore, he suggests that the parameters of the above expression should be fitted using two iterations of the procedure.

After considerable discussion concerning the roots of the characteristic equation, the author gives an example concerning the interpolation of functions of the above type on the life table published in the November 1957 issue of *Bolletino Mensile di Statistica*, Rome. This set of demographic data comes from the Italian Census of Population, 1951, and from the mortality of the male population during the years 1950-1953.

(V. Amato)

VARADARAJAN, V. S. (Indian Statistical Institute, Calcutta)
A remark on strong measurability—*In English*
Sankhyā (1958), **20**, 219-220 (3 refs.)

0.0 (–.–)

A mapping of a measure-space into a metric space is called measurable if the inverse image of every Borel set of the metric space is a measurable set. A measurable mapping is called strongly measurable if it is the almost everywhere limit of a sequence of measurable mappings each one of which assumes only a finite number of values. It is well known that if the metric space is separable, then every measurable mapping into it is strongly measurable. In this paper the author discusses this proposition when no separability assumptions are made on the metric space. Assuming the continuum hypothesis or, more generally, assuming only that the power of the class of all subsets of any uncountable set is greater than that of the continuum, the author demonstrates that every measurable mapping of a measure-space into a metric space is strongly measurable provided that the σ-field in the measure-space is generated by denumerably many sets.

(R. Ranga Rao)

BARTON, D. E. & DAVID, F. N. (University College, London) 1.3 (1.4)
Runs in a Ring—*In English*
Biometrika (1958) **45**, 572-578 (5 refs., 6 tables)

It is supposed that there are r identical beads of k colours, r_i, being the number of beads of the ith colour. These beads are supposed to be threaded on a string to form a bracelet which may not be turned over although it may be rotated. The authors point out that two different problems are involved. First, the beads may be imagined threaded on a string and the string then tied to form a ring and second, the beads may be imagined placed in a circle in which symmetries have to be allowed for. Classical writers such as Whitworth discuss the first problem and assume that it is the second.

The first problem of the line bent to the circle follows previous work on the alternations of colour in a line. Tables are given showing the exact number of alternatives for two, three, and four colours and up to twelve beads. The moments of these distributions are noted, as is also the fact that the positive binomial with the correct first two moments is a suitable approximating function.

The alternations in the second problem are given the name of Jablonski runs after the mathematician who, among others, gave a partial solution for the total number of arrangements of this kind. Two different methods of enumerating these alternations are given and a further set of tables has been calculated for those cases in which the number of Jablonski runs is different from those of the first problem.

(F. N. David)

BASU, D. (Indian Statistical Institute, Calcutta) 1.0 (4.1)
On statistics independent of sufficient statistics—*In English*
Sankhyā (1958) **20**, 223-226 (2 refs.)

The author in an earlier paper [*Sankhyā* **15**, 377] stated that any statistics independent of a sufficient statistic must have the same distribution for all values of the unknown parameter. In this paper the author gives an example to show that this proposition is not true in general. Conditions under which this is true are discussed. For this, the author needs the concept of " connectedness " between a pair of parameter points which he introduces as follows : $\theta \in \Omega$ and $\theta' \in \Omega$ are said to be connected (by a statistic T) if there exists a finite number of parameter points $\theta_1, \theta_2, \ldots, \theta_k$ which along with θ and θ' form a chain :

$\theta_0, \theta_1, \ldots, \theta_k, \theta_{k+1}$ (putting $\theta = \theta_0$ and $\theta' = \theta_{k+1}$)

such that the distributions of T overlap for any two consecutive members of this chain. Two distributions are said to overlap if there is some interval to which they assign probabilities less than unity.

The author shows that if all pairs of parameter points in Ω are connected by the sufficient statistic, then any statistic independent of the sufficient statistic has the same distribution for all values of the unknown parameter.

(S. K. Mitra)

BAYES, T. 1.0 (11.9)

Essay towards solving a problem in the doctrine of chances—*In English*
Biometrika (1958) **45**, 296-315
[Reproduced from *Phil. Trans.* (1736) **53**, 370-418]

This is a reproduction of the famous paper of the Rev. Thomas Bayes, F.R.S. It is in the form of a letter from Richard Price to John Canton, and begins, " I now send you an essay which I have found among the papers of our deceased friend Mr. Bayes, and which, in my opinion, has great merit, and will deserve to be preserved ". Price explains what he thinks Bayes has done but is not entirely explicit about his own additions and subtractions. Bayes begins with " Problem : Given. The number of times in which an unknown event has happened and failed : Required. The chance that the probability of it happening in a single trial lies somewhere between the two degrees of probability that can be named." He defines his terms and in Section I gives a series of seven propositions regarding the addition and multiplication of probabilities

Section II is devoted to the solution of the Problem. Balls are thrown randomly on a square table and the probability of their resting place is discussed. The argument leads to three rules. " Rule 1. If nothing is known concerning an event but that it has happened p times and failed q, in $p+q$ or n trials, and from hence I guess that the probability of its happening lies somewhere between any two degrees of probability as X and x, the chance that I am right in my guess is . . ." Rules 2 and 3 are variants of Rule 1. " The first rule gives a direct and perfect solution in all cases: and the two following rules are only particular methods approximating to the solution given in the first rule."

In the Appendix applications of the rule to particular problems are given. The probability of the Sun rising, a problem supposed to originate with Laplace, is discussed among others.

(F. N. David)

BRAUMANN, P. (Faculty of Sciences, Lisbon) 1.8 (0.7)

Generalisation of two theorems for multidecision statistical games—*In Spanish*
Rev. Fac. Cienc. Lisbon, 2a serie, A-Ciências Matemáticas (1959) **7**

The author deals with the problem of extending some results contained in Blackwell & Girshick's *Theory of Games and Statistical Decisions*.

He defines a symmetric multidecision problem as one with a finite number of decisions in which the integration in the sample space and the decision space may be interchanged. The paper then discusses the result that, for this kind of problem, whatever is the strategy of Nature for each second kind of strategy of the statistician their exists a first kind of strategy with the same (equivalent) risk.

Finally the author characterises some decision functions with minimum *a-posteriori* risk as invariant Bayes' strategies in some subspace of the strategies.

(J. Tiago de Oliveria)

ERDŐS, P., RÉNYI, A. & SZÜSZ, P. (Eötvös University, Budapest) **1.5 (10.1)**

On Engel's and Sylvester's series—*In Hungarian*

Ann. Univ. Sci. Budapest—Sect. Math. (1958) 7-32 (16 refs.)

The authors are dealing with metrical properties of Engel's series of the first kind and Sylvester's series (also called Engel's series of the second kind). Let the Engel's and Sylvester's series of the real number X $(0<X<1)$ be

$$X = 1/q_1+1/q_1q_2+\ldots+1/q_1q_2\ldots q_n+\ldots \quad (1)$$

and

$$X = 1/Q_1+1/Q_2+\ldots\ldots+1/Q_n+\ldots\ldots \quad (2)$$

respectively. Here q_n and Q_n are defined as follows: $q_{n+1} = \{1/r_{n+1}\}$ where $r_0 = x$, $r_{n+1} = T_1(r_n)$ and $T_1(t) = t\{1/t\}-1$ $(0<t<1)$ and $Q_{n+1} = \{1/R_n\}$, where $R_0 = x$, $R_{n+1} = T_2(R_n)$ and $T_2(t) = t-1/\{1/t\}$. Here $\{z\}$ denotes the least integer which is $\geqq z$. The following theorems are proved: (1) the probability that the digit k occurs exactly r times in the sequence q_n is equal to $k-1/k^{r-1}$. Here, and in what follows, the interval $(0, 1)$ and the Lebesgue measure is considered as the space of elementary events and the probability measure respectively.

$$\lim_{n\to+\infty} P(\log q_n-n/\sqrt{n}<y) = (1/\sqrt{2\pi})\int_{-\infty}^{y} e^{-\frac{1}{2}t^2}dt \quad (2)$$

$$P\Big(\lim_{n\to+\infty} \sqrt[n]{q_n} = e\Big) = 1 \quad (3)$$

$$P\Big(\lim_{n\to+\infty} \log q_n-n/\sqrt{2n\log\log n} = 1\Big) = 1 \quad (4)$$

$$\lim_{n\to+\infty} P\,[\log\{(Q_n/Q_1Q_2\ldots Q_{n-1}) -n\}/\sqrt{n}\}<y] = (1/\sqrt{2\pi})\int_{-\infty}^{y} e^{-\frac{1}{2}u^2}du \quad (5)$$

$$\lim_{n\to+\infty} \log Q_n/2^n \quad (6)$$

exists and is finite and positive for almost all X.

Theorems 2, 3 and 4 have been found earlier by P. Lévy and É. Borel [É. Borel, Sur les développements unitaires normaux, *Comptes Rendus* (1947) **225**, 51, É. Borel, Sur les développements unitaires normaux, *Ann. Soc. Polonaise de Math.* (1948) **21**, 74-79, P. Lévy, Remarques sur un théoreme de M. Émile Borel, *Comptes Rendus* (1947) **225**, 918-919]. The authors give new and detailed proofs for these known results—as well as for their new results. The proofs of the theorems on Engel's series are based on the fact that the sequence of random variables $q_n(x)$ forms a homogeneous Markoff chain and the random variables $\log q_n/q_{n-1}$ are almost independent.

(P. Révész)

13

FRANCKX, E. (Ecole Royale Militaire, Bruxelles) **1.6 (-.-)**

Strong Law of Large Numbers for uniformly limited variables—*In Spanish*

Trab. Estadist. (1958) **9**, 111-115

The strong law of large numbers holds for a sequence X_i of aleatory variables if the difference between the arithmetic mean S_n of the first n variables and the mathematical expectation of this mean converges to zero with probability $= 1$, when n goes to infinity. The characteristic sub-set is defined as a sub-sequence S_{n_i} such that both the Lim $u_{i+1}/u_i = 1$ and convergence to zero hold when the centres of the variables are at the mathematical expectation.

If any assumptions on the dependence or independence of the variables are made, considering only bounded variables, the following criterion is established: the strong law of large numbers holds for the uniformly limited variables if the sequence S_n has a characteristic sub-set. The variables X_i are uniformly limited if the absolute value of the difference between X_i and its mathematical expectation m_x is smaller than a certain number L which is independent of i. It can be demonstrated that this condition is necessary by the fact that the set in which S_{n_i} does not converge is contained in the set in which S_n does not converge.

It is not so easy to show that the condition is sufficient. For this the author enunciates three previous lemmas: *Lemma A.* If λ_n is a sub-sequence of positive numbers uniformly limited by L, and n_i is an increasing sequence of positive integers such that Lim $u_{i+1}/u_i = 1$ then $Sn-S_{n_i}$ becomes as small as possible choosing i sufficiently large. S_n denotes the arithmetic mean of the n first numbers of the sequence λ_n (n must satisfy the condition $n_i\geqq n<n_{i+1}$). *Lemma B.* A similar condition is established when the numbers λ_u are replaced by uniformly limited aleatory variables X_n-m_{sn}. *Lemma C.* It is proved by means of lemma B that it is equivalent to asserting that the event which states that the variable S_n-m_{sn} is greater than the absolute value of ϵ for some $k>n$ is incompatible with the event which expresses the absolute value of ϵ/i for all $n_i>n$, always provided that n be sufficiently large. Hence, it is easy to assume that the existence of the characteristic sub-set is sufficient. In fact, from the condition that the probability of convergence of $S_{n_i}-m_{S_{n_i}}$ be equal to 1, it follows that the probability of divergence of $S_k-m_{S_k}$ must be equal to zero. So the strong law of large numbers must hold.

(P. Zoroa)

14

MOTOO, M. (Institute of Statistical Mathematics, Tokyo) **1.5 (10.1)**

Proof of the Law of Iterated Logarithm through diffusion equation—*In English*

Ann. Inst. Statist. Math., Tokyo (1958) **10**, 21-28 (5 refs.)

The aim of this paper is to prove by a unified method a group of theorems of the iterated logarithm type for the absolute value of the k-dimensional Brownian motion ($k = 1, 2, \ldots$).

The author first considers a group of four classes (C) of increasing (decreasing) functions in the interval $(0, \infty)$ of a diffusion process $\{X(t)\}$. If the process is recurrent and the expectation of its recurrence time is finite—as in the Uhlenbeck process—these four classes of function are completely determined. One theorem and three lemmas are given in this section of the paper. The next section applies the theorem to the k-dimensional Brownian motion—a one-dimensional stationary diffusion process which does not satisfy the above condition of recurrence. By using a change of scale and of time it is shown how the theorem of the previous section can be rendered applicable to this case. Two theorems on the law of the iterated logarithm are proved and the paper concludes with a brief application to a general process subject to certain regularity conditions.

(C. Hayashi)

NASR, S. K. (Faculty of Science-Moharem Bey, Math. Dept., Alexandria) **1.5 (1.6)**

A law of large numbers for abstract random variables—*In English*

Metrika (1958) **1**, 89-98 (15 refs.)

The author considers random variables with values in an abstract set X in which a separating family D of pseudo-metrics is given. He recalls and discusses his definition, based on the ideas of S. Doss [*Bull. Sci. Math.* (1949) (2), **73**, 48-72], of the expectation of a random variable and the arithmetic mean Z_n of random variables $X_1, \ldots, X_n$. The sequence Z_n is called strongly stable relative to D if there is a sequence c_n in X such that $d(Z_n, c_n)$ converges stochastically to zero for every d in D. This stability is said to be normal if, in addition, the expectation of every Z_n exists and can be chosen for c_n. If X denotes a Banach space and D the family of semi-norms defined by the elements of the dual space of X, the strong normal stability of Z_n for any sequence X_n of independent random variables with a common distribution follows immediately from the classical weak law of large numbers. In comparing this result with the one of Fortet and Mourier [*C.R. Acad. Sci., Paris* (1953), **231**, 18-20] the author remarks that the latter impose restrictions on X and the X_n; however, their conclusions are, of course, much stronger, viz. almost certain strong convergence instead of stochastic weak convergence. Next the author combines his result with the one of Doss [*Publ. Inst. Statist., Paris* (1954), **3**, 135-142] asserting that in any Banach space Z_n is a sufficient estimation of the expectation within a family of normally distributed random variables with common variance. Finally, some particular cases are considered.

(K. Krickeberg)

OKAMOTO, M. (Osaka University) 1.0 (2.4)

Some inequalities relating to the partial sum of Binomial probabilities—*In English*
Ann. Inst. Statist. Math., Tokyo (1958) **10**, 29-35 (5 refs.)

This paper gives a simplified proof for a somewhat stronger inequality than that given by J. V. Uspensky, at page 102 of his book, *Introduction to Mathematical Probability* (New York, 1937) relating to the partial sum of binomial probabilities. If x is a random variable following a binomial distribution $B(n, p)$, Uspensky states that

$$P\{|\, x/n-p\,| \geqq c\} < 2 \exp(-nc^2/2)$$

where $c>0$ and $0<p<1$.

After considering two preliminary lemmas, four theorems are stated of the form $P(x/n-p \geqq c) < \exp(-2nc^2)$ (Theorem 1(i)). Theorem 1(ii) reverses the inequality within the bracket and states c with a negative sign. Theorems 2(i) and (ii) modify the exponent to $\exp(-nc^2/2pq)$ where $p \geqq 0{\cdot}5$ and $p \leqq 0{\cdot}5$ for the negative value of c. Theorems 3 and 4 take the root of x/n and p in the basic inequality with the exponent unchanged in Theorem 3 but halved in Theorem 4, which deals with the negative value of c in a similar way to Theorems 1(ii) and 2(ii).

The paper concludes with an application of Theorems 3 and 4 to Matusita's multinomial distance—a test of fit for an empirical multinomial distribution [*Ann. Math. Statist.*, **26**, 641-640 and *Ann. Inst. Statist. Math.* **7**, 67-80].

(C. Hayashi)

17

RÉNYI, A. (Math. Inst., Hungarian Academy of Science, Budapest) 1.4 (11.7)

On a one-dimensional problem concerning random place filling—*In Hungarian*
Mag. Tud. Akad. Kut. Mat. Intézet. Közleményei (1958) 109-127 (7 refs.)

The following problem is treated : how many unit intervals can be placed on the interval $(0, x)$ in the following manner : the first unit interval is placed at random on the interval $(0, x)$ (this means that its centre I is uniformly distributed in the interval $(\frac{1}{2}, x-\frac{1}{2})$; similarly, a second unit interval is chosen at random (in the same sense) independently from the first. If the second interval intersects the first it is discarded and the choice repeated till the chosen interval does not intersect the first one. The process is repeated in the same way, i.e. if there have been chosen already n (disjoint) unit intervals, we choose another interval at random, but retain it only if it does not intersect any of the previous intervals ; otherwise the choice is repeated, etc. The process is finished when there is no possibility of placing a further unit interval without intersecting any of the previous intervals. The number ν_x of the unit intervals thus obtained is a random variable ; this is studied in this paper.

The author derives recursive formulae for the moments and for the distribution function of ν_x. Introducing Laplace-transforms, and applying a Tauberian theorem, he deduces that, denoting by $M(\nu_x)$ the mean value of ν_x,

$$\lim_{x\to\infty} M(\nu_x)/x = C = \int_0^\infty \exp\left(-2\int_0^t 1-e^{-u}/n)du \sim 0.748,$$

and that ν_x/x tends stochastically to C for $x \to \infty$.

The solution of the problem has the following practical application, remarked by N. G. de Bruijn : there is a parking place consisting of a pavement of L metres length. Every car arriving places itself randomly in the intervals being still free (it does not strive to close up). The problem is how many cars can park if each one requires a free interval of length l ? It follows that with a high probability the number of these cars will be approximately $0.748\ L/l$.

(K. Sarkadi)

18

RÉNYI, A. (Math. Inst., Hungarian Academy of Science, Budapest) 1.4 (10.8)
Investigation by means of probability theory of the conductivity of certain resistances—*In Hungarian*
Mag. Tud. Akad. Kut. Mat. Intézet. Közleményei (1958) 247-256 (3 refs., 1 fig.)

Account is given of some probabilistic considerations concerning the conductivity of resistances which are manufactured by pressing the mixture of carbon-powder with the powder of some non-conducting material (e.g. phenoplast). Such resistances are, as it is well known, used in radiotechnics. The propagation of the current in the resistance mentioned above is considered as a chain reaction. The calculations have been carried out under the supposition that the carbon-powder is distributed with uniform density (according to Poisson's law) in the whole body of the resistances. The numerical results were in striking contradiction with the experimental facts. This has led Mr J. Katona (of the Industrial Research Institute for Telecommunication) to investigate the microstructure of the mentioned resistances. The microscopic pictures (one of which is reproduced in the paper) have thrown light on the reason of the contradiction between theory and facts, as they have shown that the carbon-powder is not uniformly distributed, but is contained in thin layers. This observation gave a possibility to improve the quality of the mentioned resistances.

(Author's summary)

THIONET, P. (Inst. Nat. Statist., et Études Econ., Paris) 1.1 (10.0)
On the average interval between two consecutive Friday's 13th—*In French*
J. Soc. Statist. Paris (1958) **99**, 284-293

This paper is an amusing study posing (without resolving; the problem is unsolvable) the question of applying the calculus of probabilities to the study of an isolated event. The author refers to the central point of Borel's views on this subject—that each case should be considered on its merits and without preconceived ideas.

After references to groups of equally possible events and that the observations are not always a realisation of a stochastic process, the author states that the case to be considered is not a chain process or especially ergodic. He reassures the reader as to all this mathematical vocabulary: the paper is concerned with the appearance of Friday, 13th on the calendar. It is not the approach of a new year, and the arrival of calendars, that raises this problem afresh but a paper given by Nemchinov at the ISI Conference in Brussels (September 1958). The author makes use of an application of arithmetically ordered probabilities and the following sentence summarises his point of view:

> "It seems to us that, in certain cases, events determined rigorously but by a sufficiently complex calculation so as to reveal the nature of the simple probability consideration, permitted at least a partial explanation."

The paper then deals with the statistical regularities observed in the data and the author concludes that the sequence of events includes (*a*) statistical regularities which can be explained by a simple probability rule, (*b*) statistical regularities which are unexplainable and (*c*) a tendency to irregular oscillations which also need explaining. After a section on the problem of a causal explanation, the author concludes with a critical commentary on the explanation based upon probability considerations.

(D. Dugué)

VARADARAJAN, V. S. (Indian Statistical Institute, Calcutta) 1.5 (–.–)
A useful convergence theorem—*In English*
Sankhyā (1958) **20**, 221-222 (2 refs.)

This paper discusses the convergences of a sequence X_n of k-dimensional vectors. It is well known that in order that this sequence may converge to a given vector X in distribution, it is necessary and sufficient that for every linear function l, the sequence of random variables $l(X_n)$ converge in distribution to the random variable $l(X)$. In this paper the author considers the case when the existence of X is not known beforehand, and proves that in order that there may exist a vector X towards which the X_n are convergent in distribution, it is necessary and sufficient that the distributions of the random variables $l(X_n)$ converge to some distribution for every linear function l. As a corollary, the author deduces that in order that a sequence of vectors be asymptotically normal, it is necessary and sufficient that every linear function of the vectors be asymptotically normal.

(R. Ranga Rao)

ZOROA TEROL, P. (Instituto de Investigaciones Estadísticas, Madrid) 1.8 (10.1)
A problem on sequential decision in economics—*In Spanish*
Trab. Estadist. (1958) **9**, 103-110

A trader has to sell some goods in one of n successive opportunities that can be called years 1, 2, ... n. The profit he can get in the year i is a random variable represented by x_i : the joint distribution of $(x_1, \ldots, x_n)$ is known by assumption. The trader has to choose the year for selling, considering the profit he may get for this year and the precedent ones. Let a be the profit obtained when m non-selling years had passed. Then the sale must be performed if this year is reached without selling and $x_m \geqslant a$. So the variable x_m is changed into another one X_m which represents the best profit the trader can get if the year m is reached without selling. Its value is x_m when $x_m \geqslant a$ and a when $x_m < a$.

The best profit X_{m-1} obtained if the year $m-1$ is reached without selling can be deduced from the knowledge of X_m. Its value is either x_{m-1} when x_{m-1} is bigger than the profit expected for the year m or the value expected for the year m when x_{m-1} is smaller than the profit expected if the sale is realised during the year m. In the first case the sale must be performed in the year $m-1$ and in the second one it would be better to sell in the next year. Following on this argument a recurrent law may be found in order to establish the best strategy for selling. Another similar situation is stated assuming that the last year in which the sale is possible is not known for sure, but this case will be ruled by a random principle.

The first problem is generalised to the case of a continuum of possibilities, i.e. that a customer may arrive at any instant in an interval of time $0 \leqslant t \leqslant 1$. The arrival of the customer obeys a Poisson's process. If there is a possibility of performing the sale then the distribution of the variable x_t which represents the profit at this moment is known. The condition satisfied by the solution of this problem is found and is applied to a simple case.

(S. Rios)

CROW, E. L. (Nat. Bureau of Standards, Boulder, California) 2.5 (2.1)
The mean deviation of the Poisson-distribution—*In English*
Biometrika (1958) **45**, 556-559 (4 refs., 1 fig.)

After deriving the mean deviation of the Poisson distribution, the author considers the behaviour of the ratio $R(m)$ of the mean deviation to the standard deviation as the mean m increases from zero. R is shown to follow a continuous damped oscillation with infinitely many maxima and minima, the minima occurring at integral values of m and the maxima being equally spaced between the minima. The slope of R is always finite but is discontinuous at integral values of m.

The value of R for a normal distribution is $(2/\Pi)^{\frac{1}{2}}$ and this is attained by the Poisson R once and only once between each adjacent pair of maximum and minimum values. It is shown that R is a poor criterion for distinguishing between the Poisson and normal distributions, or for determining whether a Poisson distribution is approximated by the normal.

An asymptotic expression for $R(m)$ is obtained and this gives the pair of points between successive minima at which R takes the normal value. The accuracy of the asymptotic expression is discussed with reference to the maximum and minimum values it gives for $R(m)$ as m varies, and also with reference to the asymptotic ratio of the maximum positive difference from the normal R, to the average of the absolute values of the adjacent maximum negative differences.

The ratio R for the binomial distribution is also considered, using an asymptotic formula given by Johnson (*Biometrika*, 1957) in his note on the mean deviation of the binomial distribution.

The behaviour of $R(np)$ is similar to that of the Poisson $R(m)$, and when np is varied by varying p with n fixed, analagous results are obtained approximately. Since these are confirmed for small n by exact calculations, it is concluded in particular that for any fixed n there are exactly $2n$ values of p for which the binomial R takes the normal value.

For fixed p and varying n, R is not continuous but its behaviour can be investigated by regarding n as continuous. Some points of this R are plotted in a diagram for two values of p. This diagram also shows the behaviour of the binomial ratio for two fixed values of n, and the Poisson and normal ratios for m between 0 and 10. The upper bound of R for any distribution is attained with p, n having values $\frac{1}{2}$ and 1 respectively.

(D. Allibone)

DRONKERS, J. J. (The Hague) 2.6 (3.8)
Approximate formulae for the statistical distributions of extreme values—*In English*
Biometrika (1958) **45**, 447-470 (8 refs., 1 table, 3 figs.)

A sample of size n is supposed with the observations ranked in order of magnitude, say $x_1, <x_2, <\ldots, <x_n$. Approximate formulae are derived for the distribution of the mth value, i.e. the mth from the right. The limiting distributions of extreme values, given by Fisher-Tippett and Gumbel, are well known. The author generalises the extreme value techniques to get his approximate formulae. Three types of functional form are supposed for the initial distribution of any x_j, viz. : exponential, Cauchy, and finite range. The basic conditions required to be imposed on these initial distributions are discussed. A formula is obtained for what is designated " the excess " i.e. the difference between unity and the cumulative distribution function of any x.

The general form of the distribution of the mth largest x is then considered. Formulae are derived whereby the mode and the maximum value of this distribution can be determined. The behaviour of the maximum value for varying n and m is studied and approximate formulae for the distribution of the mth largest x are deduced. The range of application of each formula is delineated. The limiting formulae for mth largest x for the three types of initial distribution are found and the well-known limiting extreme-value distributions are also achieved.

(F. N. David)

FERRERI, C. (School of Statistics, Palermo) **2.5 (1.1)**

An important new model of dependence—*In Italian*

Ann. Fac. Econ. Com. Palermo (1958) **12**, 221-238 (6 refs.)

The author has constructed an important theoretical model founded on the simple assumption of a constant dependency of successes: the unfavourable outcomes are supposed equally probably but without effect. The mathematical properties of the model are explained by a series of drawings of black and white balls. The author forms the one-sided dependency by replacing $1+\Delta$ white balls for every white ball that is drawn or only the drawn ball when the ball is black. He has found, for the case of n (the number of trials) very large and rare events with slight dependency, the probability distribution function P_x of x successes in n trials

$$P_x = [e^{-\lambda}/\Gamma(\lambda/d)]\,[(1-e^{-d})^x\,\Gamma(x+\lambda/d)/x\,!]$$

where $d = \log(\sigma^2/M)$ is the dependency parameter, and $\lambda = np_x$ (p_x being the probability of drawing a white ball at the nth trial when x successes have been obtained in the preceding $n-1$ trials).

The new model is a generalisation of the Poisson, negative binomial and Polya distributions. The parameter λ is related to the variance σ^2 and to the mean M by the relation $d = M^2d/2-M$.

(F. Galantino)

25

JOHNSON, N. L. (University College, London) **2.7 (4.3)**

The mean deviation, with special reference to samples from a Pearson Type III population—*In English*

Biometrika (1958) **45**, 478-483 (1 ref., 6 tables)

A random sample of n independent observations from a Type III distribution with a single parameter, α, is, at first, considered. The expected value of the sample mean deviation, m_n (written m in the text), is found after making certain linear transformations on the observed variables: for larger n, two approximations are suggested. The population value of the mean deviation is the limiting value of the expected value of the sample mean deviation as $n\to\infty$. A table is provided showing the exact values of the ratio $\mathscr{E}(m_n/\sigma)$, σ ($=\sqrt{\alpha}$, for the Type III population) being the standard deviation of the distribution, for different n and $\alpha = 1(1)8, \infty$. A comparison of the values for different α shows that, for $n \geqq 15$, a small error in α will not have any serious effect on the unbiased estimator, $m_n[\sigma/\mathscr{E}(m_n)]$ of σ. A second table, calculated from the second approximation—which is better than the first, shows the values of $\mathscr{E}(m_n/\sigma)$ thus calculated are in good agreement with the exact values, especially for $n \geqq 10$.

The relative efficiency of two estimators of σ, $m_n[\sigma/\mathscr{E}(m_n)]$ and S, the sample standard deviation is discussed in general for any parent population. This is achieved by making certain large sample approximations to the expectations and variances of m_n and S; the adequacy of the approximations to the variance is investigated for the particular case when the parent population is Normal. A comparison of approximate coefficients of variation of m_n and S leads to the conclusion that $m_n[\sigma/\mathscr{E}(m_n)]$ will have a smaller standard error than S if $\beta_2 > -3+4\sigma^2/[\mathscr{E}(m_n)]^2$, where β_2 is the second moment ratio of the distribution of the observed variable. Specific applications of the above result are made to Type III, VII and II populations, each having a single parameter, denoted by α. Three tables, one for each Type, are produced showing values of β_2 and $4\sigma^2/[\mathscr{E}(m_n)]^2-3$, for different α. It is concluded that sample mean deviation may, to the degree of approximation used, provide an unbiased estimator of σ, with a smaller standard error than sample standard deviation, even for moderately leptokurtic populations. A modified form of the above criterion is suggested for finite n.

(B. K. Ghosh)

26

KAMAT, A. R. (Fergusson College, Poona)
Hypergeometric expansion for incomplete moments of the Bivariate Normal Distribution—*In English*
Sankhyā (1958) **20**, 317-320 (2 refs., 2 tables)

2.3 (2.1)

This paper can be regarded as a continuation of the author's previous paper in *Biometrika* (1953) **40**, 20-34. In this new paper, the author expresses the incomplete moment $[m, n]$ of the standardised trivariate normal distribution in terms of the well-known hypergeometric series. The incomplete moment $[m, n]$ is defined to be the contribution to the expectation of $x^m y^n$ from the positive quadrant. This is a simple multiple of the m, nth moment of the distribution when the variables are confined to the positive quadrant. The expressions for the absolute moments are at once deduced from those of the incomplete moments. The author has tabulated the incomplete moment $[m, n]$ for all integers m, n such that $m+n \leqq 4$ and $p = -1\,(0{\cdot}1)+1$.

(J. Sethuraman)

KAMAT, A. R. (Fergusson College, Poona)
Incomplete moments of the trivariate normal distribution—*In English*
Sankhyā (1958) **20**, 321-322 (1 ref.)

2.3 (2.1)

In this paper, which may be regarded as a sequel to the author's other paper in this issue of *Sankhyā*—and both being in the nature of a continuation of the author's previous paper in *Biometrika* (1953) **40**, 20-34, the author defines the incomplete moment $[l, m, n]$ for the standardised trivariate normal distribution as the contribution from the positive octant to the expectation of $x^l y^m z^n$. Exact expressions of $[l, m, n]$, when l, m and n are integers such that $l+m+n \leqq 4$, are given without proof. For the method of derivation the reader is referred to the author's paper in *Biometrika* (1953).

(J. Sethuraman)

KUDO, A. (Kyushu Univ., Japan, & Indian Statist. Inst., Calcutta) 2.3 (5.5)
On the distribution of the maximum value of an equally correlated sample from a normal population—*In English*
Sankhyā (1958) **20**, 309-316 (5 refs., 2 tables)

In this paper a series expansion in terms of successive derivatives of the standardised normal distribution function is given for the probability integral of a multivariate normal distribution in which each variable has the same mean zero, the same variance σ^2 and the same correlation ρ with every other variable. For the distribution of the maximum of an equally correlated sample from a normal population, what is required is the above integral with all limits of integration equal. In this case the series expansion can be more simply expressed in terms of certain functions $A_{N,n}(x)$ which are defined by means of recurrence relations involving derivatives of the standardised normal distribution function. $A_{N,n}(x)$ is tabulated for $N = 1(1)10$, $n = 0(2)8$, $x = 1{\cdot}8(\cdot1)3{\cdot}5$.

In the final section the author mentions the following application: $x_1, x_2, \ldots, x_{N_1+N_2}$ are N_1+N_2 observations from a normal population with unknown mean m and known variance σ^2. Let x_M be the maximum of $x_1, \ldots, x_{N_1}$ and $\bar{x}$ the mean of $x_1, \ldots, x_{N_1+N_2}$. The author has shown elsewhere [*Sankhyā* **17**, 67] that the statistic $(x_M-\bar{x})/\sigma$ is optimum for testing the significance of an outlying observation provided it is known *a priori* that there may be just one x_i among $x_1, \ldots, x_{N_1}$ which is from another normal population with the same variance but with a greater mean value. As $y_j = x_j-\bar{x}\,(j = 1, \ldots, N_1)$ are equally correlated, the previously developed theory gives the necessary distribution.

(S. John)

PARSONS, P. A. (Dept. of Genetics & 'Trinity College, Cambridge) 2.4 (1.2)
Equilibria in auto-tetraploids under natural selection for a simplified model of viabilities—*In English*
Biometrics (1959) **15**, 20-29 (2 refs., 4 tables)

It is pointed out that the study of gene frequencies for disomic populations has an immediate analogy with that for tetrasomic populations when the latter is considered in terms of frequencies of the gametic genotypes. The resulting recurrence relations are established before and after selection when there are two alleles present.

The author states that the "determination of equilibrium points for five different viabilities is difficult" and the rest of the paper is concerned with the simplified model mentioned in the title. Verbally, this model is characterised by the fact that selection is symmetric for the two alleles, i.e. is determined by the "degree" of heterozygosity of the organism at that locus. By standardisation of the viability of the most heterozygous (G_2g_2) organism to unity the number of parameters is thereby reduced from five to two.

This situation is then examined under selfing for:

(i) Equilibria
(ii) Stability of these equilibria.

The two trivial equilibria of homozygosity are noted: these are both stable if the viability of the homozygous (G_4) organism is greater than $\frac{1}{3}$ and also than that of the (G_3g) organism.

There is always a symmetric equilibrium whose frequencies the author expresses in terms of the heterozygous gamete frequency. This frequency is tabulated for viabilities [0, $\frac{1}{4}$ ($X2$) 256] of the heterozygous organisms.

There may be a pair of asymmetric equilibria, the conditions are established and the equilibrium frequencies are tabulated (where existent) for the same viabilities as above.

The method of determining stability is displayed, the results are tabulated and discussed briefly.

(C. W. Clunies-Ross)

RAMASUBBAN, T. A. (London School of Economics) 2.1 (2.5)

The mean difference and the mean deviation of some discontinuous distributions—*In English Biometrika* (1958) **45**, 549-556 (1 ref., 5 tables)

The author considers the following discontinuous distributions: the negative and positive binomial, Poisson, the logarithmic and the geometric. For these distributions he finds algebraic expressions for the mean difference, Δ. For the hypergeometric, the negative binomial, Poisson, the logarithmic and the geometric distributions he finds algebraic expressions for δ, the mean deviation about the mean. The author shows that the usual definitions of the mean difference and the mean deviation (about the mean) can be considered as particular cases of more generalised forms. Furthermore, associated with these more generalised measures are two ratios involving the variance of the distribution as a standardising factor: one of these ratios is an extension of Geary's ratio [*J. R. Statist. Soc.* (1930), **93**, 442]. No question of sampling is involved and a further paper is envisaged extending this study.

(F. N. David)

STUART, A. (London School of Economics) 2.3 (6.6)

Equally correlated variables and the multinormal integral—*In English J. R. Statist. Soc.* B (1958) **20**, 373-378 (8 refs.)

The joint distribution of n random variables all equally positively correlated and with correlation equal to ρ, can be constructed from a set of $n+1$ random variables. The following result is demonstrated. "Any statement concerning the means of n equally correlated variables with arbitrary means, equal variances, σ^2, and correlations $\rho, (-1/(n-1) \leqq \rho < 1)$, is precisely equivalent to the same statement concerning n uncorrelated variables with the same means and equal variances $\sigma^2(1-\rho)$. Any statement concerning σ^2 for the equally correlated variables is directly translatable into one concerning $\sigma^2(1-\rho)$ for the uncorrelated variables." The result is applied (i) to obtain a t test for equally correlated variables (ii) to write down an integral for the evaluation of the positive secant of the multivariate normal surface where the variables are equally correlated (iii) to remark on the fact that problems regarding randomisation sets can be reduced to problems regarding independent normal variables. Equally negatively correlated variables are also discussed.

(F. N. David)

TALACKO, J. (Marquette University, Milwaukee) **2.9 (2.1)**

A note about a family of Perks' distributions—*In English*
Sankhyā (1958) **20**, 323-328 (4 refs., 1 table)

It is well known in the classical probability theory that there exist bell-shaped distributions other than the symmetric distributions like Gauss', Student's, Cauchy, etc. This paper is concerned with a class of such distributions called " Perks' functions ". The original Perks' function is a ratio of two exponential polynomials. The case when the density function is of the special type $c/(e^x+k+e^{-x})$ where c and k are parameters, is discussed in a detailed manner. For $k = 0$ and 2 this becomes the density of a hyperbolic cosine and logistic respectively. The characteristic function of this density is found for all values of k and compared with the characteristic functions of Normal, double exponential and Cauchy distributions. The moments are found in terms of the Euler and Bernoulli numbers and the first four moments are calculated and listed in a table for the Normal, logistic, hyperbolic secant and double exponential densities.

Finally, linear combinations of independent random variables with a double exponential distribution are considered and it is shown that by properly choosing the combination we can arrive at the Normal, degenerate and logistic distributions. Further, a counter-example for the central limit theorem is constructed by the same method. If instead of the double exponential distribution we start with a hyperbolic secant, it is pointed out that similar relationships can be established.

(K. R. Parthasarathy)

WHITTLE, P. (Applied Maths. Lab., D.S.I.R., New Zealand) **2.1 (0.2)**

On the smoothing of probability density functions—*In English*
J. R. Statist. Soc. B (1958) **20**, 334-343 (6 refs.)

Extending the method of a previous paper [P. Whittle, *J. R. Statist. Soc.* B (1957) **19**, 38-47] on the smoothing of frequency spectra, the author considers the estimation of a probability density function by means of a linear smoothing of an observed (sample) density. Thus, the estimate (at each value of the variate x) is found by evaluating a certain weight function (depending on x) at each of the sample values and averaging over the sample. The weight functions are to be determined by some optimality criterion.

Some sort of model must be postulated if there is to be any basis for estimation. The author proceeds to consider the curve being estimated as one of population of curves and for which population the correlation between adjacent ordinates approaches unity. This does not prevent any one curve from having many kinds of discontinuity and irregularity.

The weight functions are then chosen to give unbiased estimates with minimum variance (relative to this population). The problem reduces to that of solving a certain integral equation. It is shown that the resulting estimates are invariant with respect to monotone transformations of the variate scale.

In the case where, in the population, the correlation between different ordinates depends only on the difference between the abscissae, the integral equation can be solved by Fourier transforms : at least for the case of a distribution on the whole real line. A particular example is solved explicitly : for this example, it is shown that the effect of considering distributions over only a finite interval is negligible for large sample except at the ends of the interval.

The asymptotic behaviour of the mean square deviation is examined and is shown in general to depend on the behaviour of the autocovariance function for small displacements. The details for the general case depend on the density near zero of the eigenvalues of the kernel in the integral equation for the weight function.

(C. L. Mallows)

DARWIN, J. H. (App. Maths. Lab., D.S.I.R., New Zealand) 3.1 (0.4)
On corrections to the Chi-squared distribution—*In English*
J. R. Statist. Soc. B. (1958) **20**, 387-392 (4 refs., 1 table)

The null hypothesis distribution of the Chi-squared goodness-of-fit test function (here called χ^2) is considered when the null hypothesis is fully specified for a multinomial distribution. When the number of observations (N) is large the distribution of χ^2 is approximately that of χ^2 with $k-1$ degrees of freedom. (K. Pearson, *Phil. Mag.*, 1900).

An asymptotic series of correction terms has been found by Hoel (*Ann. Math. Statist.*, 1938) where these terms form a sequence of continuous varying functions whose coefficients are of increasing power in $1/N$. It is also known that when N is small the discrete nature of the variation of χ^2 renders such asymptotic expansion inadequate. It is this effect which is treated in the present paper. For large N as k increases, the jumps in the cumulative probability distribution function of χ^2 are shown to be of order $1/N^{k/4}$, and an expression is obtained for their magnitude.

The method used is to express the multinomial distribution in terms of the Fourier inversion of the joint characteristic function of the multinomial variables. This is used to obtain the Laplace transformation of the probability density function of χ^2 which can be expressed in terms of theta functions. The properties of these functions are used to express the Laplace transformation in such a way that it can be split into two parts ; one involving Hoel's corrections and the other, the major term, retaining the discrete nature of its variation. This latter is expressed as a multiple series of functions which are first modified by neglecting a set of terms of order $1/e^N$ and then Laplace-inverted in terms of Bessel functions of large argument, $O(N^{\frac{1}{2}})$ which are replaced by their asymptotic expansions. The result is an expression which is shown to be of order $1/N^{k/4}$.

The analysis is checked and illustrated by considering the binomial case, $k=2$, where the probability distribution function has only a simple series of discontinuities. The principal term in the expression for these is evaluated. A numerical illustration is given for the case $N=10$, $p_1=2/7$ where the cumulative probability distribution function of χ^2 is considered at the jump points together with the approximation given by the cumulative probability distribution function of χ^2 with one degree of freedom together with two corrected expressions (using the author's results). The two forms of correction arise since the correction term is not only a step function taking different values either side of a jump point but also takes an intermediate value at the jump point. This allows the cumulative probability distribution function of χ^2 (a step function) to be approximated by the cumulative probability distribution function of X^2 with one degree of freedom evaluated either at the lower edge of the step with the correction added, or at the upper edge with the correction subtracted. In the case considered the author's corrections give a considerable improvement.

(D. E. Barton)

35

GEFFROY, J. (Caen University) 3.8 (1.6)
A contribution to the theory of Extreme Values—*In French*
Publ. Inst. Statist., Paris (1958) **7**, 37-121 (3 figs.)

This paper is the first part of the thesis presented by the author for his Doctor of Science.

The problem which sets the foundation for all this research is the study of stochastic behaviour of the " middle " of a sample. That is to say, half the sum of the greatest and smallest values, i.e. the mid-range. The author supposes all the variables to be of a single dimension. The first chapter is devoted to the over-estimation and under-estimation of extreme values according to the three topologies used in this work : convergence in probability, almost certain convergence, almost completely certain convergence.

In the second chapter the stability of extremes in the sense of Kolmogoroff is approached. The author in this way rediscovers some theorems of which certain are due to Gnedenko : the examples use the concentration function of Paul Lévy.

The last two chapters are used to give the conditions of almost completely certain convergence and almost certain convergence (this latter case being the more difficult). The author has here concerned himself with new results. This paper will be followed by another which will be published in a subsequent number of this same journal.

(D. Dugué)

HOOPER, J. W. (Netherlands School of Economics, Rotterdam) **3·5 (3·9)**

The sampling variance of correlation coefficients under assumptions of fixed and mixed variates—*In English*

Biometrika (1958) **45**, 471-477 (4 refs., 2 tables)

The writer considers T independent observations of the $M+\Lambda$-variate (y, x) where $y \equiv (y, \ldots y_M)$ and $x \equiv (x, \ldots x_\Lambda)$ from the same population. He contrasts the models (i) where (y, x) is a multinormal variate with zero mean, (ii) the regressive model where x is a parameter (" fixed variate ") and where $Y = \mathscr{E}(y) = x\pi$ (where π is an $M \times \Lambda$ matrix) and $y - Y$ is an M-variate multinormal. He proposes a " mixed " model to include (i) and (ii) as special cases (iii) where $x = \xi + w$ where ξ is a parameter and w is a Λ-variate multinormal with zero mean ; the vector $y - x\pi$ being multinormally distributed independently of w. Under (iii) the asymptotic sampling variance (as $T \to \infty$) of a sample canonical correlation coefficient (r) is given as $2T$ var $(r) = (1-\rho^2)^2(2-\rho^2p^2)$, where ρ is the corresponding parent canonical correlation and $p = \mathscr{E}(\xi \times \xi')$, the variables being in canonical form so that $\mathscr{E}(xx') = 1$, $0 \leqslant p \leqslant 1$: proof being left to an appendix. The asymptotic sampling covariance of any pair of sample canonical correlations is shown to be zero.

Three special cases are considered : $M = 1$, $\Lambda = 1$; $\Lambda > M = 1$; $w \equiv 0$ (i.e. $p = 1$) corresponding respectively to bivariate (zero order) correlation, and models (i) and (ii). Tables are given of T var (r) for $\rho^2 = 0(0{\cdot}1)1$, $p^2 = 0(0{\cdot}1)1$, and of T var $(R^2) = 2\bar{R}^2(1-\bar{R}^2)^2(2-\bar{R}^2p^2)$ for $\bar{R}^2 = 0(0{\cdot}1)1$ and $p^2 = 0(0{\cdot}1)1$ (where R is the sample multiple correlation coefficient, $M = 1$, and $\bar{R}$ its population value). The writer notes that the asymptotic sampling variances are continuous in p, $0 \leqslant p \leqslant 1$, and that the choice of models (i) $(p = 0)$ and (ii) $(p = 1)$ or (iii) makes a considerable difference in the asymptotic sampling variances of r and R (for large ρ, R) and in that of the zero order correlation (but *not* the regression) coefficient. The appendix proves the stated results by means of elementary but lengthy algebra heavily dependent on the underlying assumption of normality : matrix notation is not used.

(D. E. Barton)

JAMES, G. S. (University of Leeds, England) **3·7 (3·9)**

On moments and cumulants of systems of statistics—*In English*

Sankhyā (1958) **20**, 1-30 (16 refs.)

The author presents a direct proof of Fisher's sampling rules for k-statistics which is both reasonably rigorous and easier to follow than the earlier ones and shows how the results can be modified to obtain the sampling properties of a general set of statistics z_1, z_2, z_3, ... (homogeneous polynomial symmetric functions of the sample values, of degrees 1, 2, 3, ... respectively). The proofs are mostly based on the Fisher-Kaplan approach, univariate and less general results being obtained in special cases. Simple explanatory examples have been introduced in various places. The author also defines a series of l-statistics whose expectations are population moments about the mean and whose sampling properties can be investigated by using those of z_r. Rules are provided for obtaining (1) the expression of (ordinary or mixed) moments of the z_r in terms of population moments, (2) the expression of (ordinary or mixed) moments of the z_r in terms of the cumulants, and (3) the expression of (ordinary or mixed) cumulants of the z_r in terms of cumulants.

The corresponding rules for the z-statistics appropriate to samples from a p-variate population are also given. Proofs of all the rules stated are given.

(P. K. Kumaran)

MERRINGTON, M. & PEARSON, E. S. (University College, London) **3.2 (2.7)**
An approximation to the distribution of non-central *t*—*In English*
Biometrika (1958) **45**, 484-491 (5 refs., 2 figs., 1 table)

The probability density distribution and first four moments of non-central t are stated. The relationship between the moment-ratios, β_1 and β_2, and the parameters f and δ is explored. A diagram is given of contours of constant f and δ in the β_1, β_2 plane. The beta points for the non-central t are found to be entirely within the area corresponding to Pearson Type IV curves. As δ increases the limiting form of the distribution is that of the reciprocal of χ.

Comparisons are made between the true upper and lower 5, 1, and 0·5 per cent. points and those found by interpolation in the " Percentage points of Pearson curves," Table 42 in *Biometrika Tables for Statisticians*. True values were found either from N. L. Johnson and B. L. Welch's paper: " Application of the non-central t-distribution," [*Biometrika*, **31**, 362-389] or from " Tables of the non-central t-distribution " by G. J. Resnikoff and G. J. Lieber-mann. At the six percentage points, the Pearson Type IV curve provides a very good approximation to the non-central t-distribution over a wide range of values of f and δ. The difference between the true and approximate percentage points being never more than one in a hundred.

Inaccuracies were found in the Resnikoff and Lieber-mann tables of percentage points for high values of f. The investigation demonstrates the practical utility of tables which enable a simple system of non-normal curves to be used in approximating to the percentage points of other untabled distributions through the use of four moments.

(M. Merrington)

SRIVASTAVA, A. B. L. (Statist. Lab., Indian Inst. of Tech., Kharagpur) **3.2 (3.1)**
Effect of non-normality on the power function of t test—*In English*
Biometrika (1958) **45**, 421-430 (12 refs., 3 tables)

The author uses techniques which have been made familiar by Gayen and Ghurye. Starting with a parent population which may be described by four terms of Edgeworth's series to obtain expressions in terms of the *Hh* functions by means of which the power function can be calculated. Some calculations are given showing the effect of non-normality on the power function for samples of sizes 5, 10, and 20 and for the standardised difference of means 0, 1, 2, 3, 4, 5. The inverse problem of the determination of the critical region corresponding to a predetermined level of significance is also considered using the Cornish-Fisher expansion. The general conclusion is that unless the parent population represented by the Edgeworth series is markedly non-normal the effect of non-morality is not marked. When kurtosis in the parent population is not very different from that in a normal population the skewness effect becomes more pronounced.

There is an appendix by E. S. Pearson in which the correspondence between the theoretical results of Mr. Srivastava and the early (1929) empirical sampling results of Pearson and Adyanthaya is pointed out.

(F. N. David)

BAHADUR, R. R. (Indian Statistical Institute, Calcutta) **4.1** (4.3)
Examples of inconsistency of maximum likelihood estimates—*In English*
Sankhyā (1958) **20**, 207-210 (4 refs.)

The maximum likelihood (M.L.) method provides an estimate of the entire population distribution even if one is interested only in a particular parameter. It is, therefore, of interest to enquire whether the maximum likelihood estimates of the population distribution converge, in some given intrinsic sense, to the actual distribution as the sample size tends to infinity. After remarking that Wald's famous proof of the consistency of the M.L. estimates can be formulated so as to yield regularity conditions which ensure convergence in this sense, the author states that the need for regularity conditions is not clear from previously published examples of the failure of the M.L. method. This is because these examples are based either on convergence definitions which are far from intrinsic or on the discontinuous nature of the parameter to be estimated, considered as a functional on the set of alternative distributions.

The author then gives two examples where maximum likelihood estimates of the population distribution function fail to converge to the actual distribution function : convergence being understood in the sense of P. Lévy. In the first example the requirement that the estimate maximise the likelihood is so inconsequential that some M.L. estimates do not converge at all. The set of alternative distributions in this example is the set of all probability distributions on [0, 1] which are absolutely continuous, having continuous densities $f(x)$ such that $0 \leqq f(x) \leqq 2$. In the second example the author constructs a countable family of admissible distributions on the set of integers, such that although M.L. estimates exist, they always converge to a fixed one from the admissible set, no matter what the actual distribution may be.

(V. S. Varadarajan)

BASU, D. (Indian Statistical Institute, Calcutta) **4.9** (8.2)
On sampling with and without replacement—*In English*
Sankhyā (1958) **20**, 287-294 (3 refs.)

The author considers a population of N identifiable units and discusses estimation procedures for the mean of some observable characteristic in this population under various types of sampling schemes—sampling with or without replacement and with equal or unequal probabilities.

In sampling with replacement and with equal probability for a fixed sample size of n the author shows that the mean of the ν distinct units in the sample (ignoring repetition of units that happen to occur more than once in the sample) is an unbiased estimate for the population mean and has uniformly a smaller variance than the sample mean. The proof is similar to that of Rao-Blackwell theorem and utilises the fact that the observations corresponding to the distinct sample units constitute a sufficient statistic in this problem which, incidentally, is not a complete one. The same result is proved in a different way by Des Raj and Khamis [*Ann. Math. Statist*, **29**, 550].

In sampling with replacement and with unequal probability, the author shows that the observations corresponding to the ν distinct units along with the associated probabilities of selection are sufficient. Hence it is possible to improve upon the sample mean by taking its conditional expectations given the values for the sufficient statistic. The actual computation of this conditional expectation is, however, very complicated, and even more troublesome is the computation of an estimate for its variance.

The author also considers sampling with replacement and with equal probability for a fixed number ν of distinct units and shows that the mean of all the n sample observations, though an unbiased estimate of the population mean, has uniformly a larger variance than the mean of the ν distinct units.

The author similarly considers sampling without replacements and with unequal probabilities and shows that the set of ordered sample observations along with their corresponding probabilities of selection are sufficient in this case. Hence any estimator which makes use of the order in which samples were drawn can be uniformly improved upon by taking its conditional expectation for given values of the sufficient statistic.

The paper also gives the sampling distribution of ν, the number of distinct units in a sample of fixed size drawn with replacement and with equal probability and gives only the first two moments of this distribution for the unequal probability case.

(S. K. Mitra)

BULMER, M. G. (Department of Social and Preventive Medicine, Manchester) **4.4 (7.1)**

Confidence intervals for distance in the analysis of variance—*In English*

Biometrika (1958) **45**, 360-369 (9 refs., 6 tables, 1 fig.)

A " parametric " model of the analysis of variance is considered, $E(Y) = X\beta$ where Y is a $n \times 1$ column vector of observations with a variance σ^2, X is an $n \times k$ matrix of known constant coefficients and β is a $k \times 1$ column vector of parameters. The null hypothesis specifies $r(\leqq k)$ parameters (or linear independent functions of these), denoted by Y_0, and $r \times 1$ column vector.

Two types of confidence interval for the " distance ", δ, a function of $(Y - Y_0)$, of the true hypothesis, Y, from the null hypothesis, are constructed : (i) a " simultaneous " confidence interval, ensuring *at least* a preassigned confidence coefficient, for the " distances " from *several* null hypotheses —which can also be used in conjunction with Scheffé's (listed in references) method of assigning simultaneous confidence intervals on individual parameters ; (ii) an approximate confidence interval narrower than (i), for this " distance " from a single null hypothesis, with a confidence coefficient *as near as possible* to some preassigned value. The methods discussed can also be used to find, in chi-square test for goodness of fit, the confidence interval for the distance between the true and null probabilities of an observation falling in a particular class.

The confidence interval (i), expressed as a simple function of the " testing " and " error " sums of squares, is illustrated by an example on analysis of variance. The idea is extended to find the confidence intervals separately, for several components into which Y can possibly be split up —an example of chi-square test for goodness of fit has elucidated such an application.

The confidence interval in (ii) is obtained by approximating this distribution of the minimum value of the " testing " sum of squares, under a particular alternative $Y = Y_1$, by a chi-square distribution with correct mean and variance. The actual computation of this approximate confidence interval, as illustrated by an example, involves mainly two quantities : (*a*) $\lambda = \delta/\sigma^2$, which can approximately be obtained by an interative process from the percentage points of F-distribution with fractional d.f.'s one of which is a function of λ, (*b*) the evaluation of the percentage points of an F-distribution for fractional d.f.'s f_1 and f_2, a special table having been provided for 1 and 5 per cent. when $f_1 \leqq 2 \cdot 2$ and $f_2 = \infty$. The adequacy of the approximation is investigated in the last section where it is found that the true confidence interval is slightly narrower than is asserted by the approximation.

(B. K. Ghosh)

CORREA POLIT, H. **4.4 (3.3)**

Statistical inference on the parameters of non-normal populations—*In Spanish*

Trab. Estadist. (1958) **9**, 117-140 (7 refs.)

A discussion about statistical inference in cases where the parent population is continuous with finite mean and finite standard deviation, but non-normal. Having in mind the possible application of results to practical situations, the author sets as a postulate that finite populations, in the case of simple random sampling, are random samples of infinite populations : " Not a random sample in the restricted sense given in the standard texts of sampling theory and methods, but rather in the more general and intuitive sense which serves as a basis of all statistical considerations."

Taking a specially defined sample of means from continuous distributions, conditioned with respect to a previous random sample, the author discusses the possibility of obtaining a sample of means of values, when these values are directly obtainable from the population.

The case is then discussed of confidence intervals when its length, or the confidence coefficient, are random variables ; the size of the sample being previously fixed. Substituting an approximate confidence coefficient for the exact one, confidence limits are set for the mean of the sampling distribution of sample means : considering both the cases of infinite and finite parent populations—using, in the latter case, the author's postulate mentioned above. Afterwards the author's attention is directed towards stratified sampling and he proceeds to extend his first postulate considering each finite stratum or sub-population as a random sample of an infinite population, in a way similar to the previous one for a non-stratified situation.

Next is discussed the case of confidence intervals when the number of the elements in the sample is a random variable, the confidence coefficient and the length of the intervals being fixed in advance, or at least an upper bound for the confidence interval. The author derives his proof from a lemma which is the well-known property that a variate t may be generated by dividing a normal variate with zero mean and unit standard deviation into the square root of an independent chi-square variate divided by its number of degrees of freedom. The two cases of infinite and finite populations are also considered and confidence intervals are set for the mean of the sampling distributions of sample means. Rules are also given for the computation of the sample size, depending on the number of samples : Special consideration being given in this case to stratified populations.

(F. Azorín P.)

DUNN, O. J. (Iowa State College ; now at University of California, Los Angeles) 4.2 (4.8)
Estimation of the means of dependent variables—*In English*
Ann. Math. Statist. (1958) **29**, 1095-1111 (7 refs., 2 tables)

Several methods are given for deriving sets of simultaneous confidence intervals for the means of a multivariate normal distribution.

Exact simultaneous confidence intervals are obtained (Section 3) for the case of known variances and for the case of equal, but unknown variances. Knowledge of the covariances is unnecessary.

Several sets of intervals having bounded confidence level are obtained (Sections 4-7) for the two aforementioned cases. In addition, the method which employs Hotelling's T-statistic is used (Section 4.2) to derive intervals for the general case where the variances are unknown and arbitrary ; here the sample size is assumed greater than the number of variables under study.

Finally, the various methods presented are compared in terms of the expected values of the lengths of resulting intervals. This is done numerically, for confidence coefficient 0·95 and several sample sizes. For general use, the winner in this respect is the method of Section 6 based on a Bonferroni inequality.

The problem of dependent variables in which means of one or more variables depend on observations (not just means) of one or more other variables, though illustrated in examples in Section 2, is not treated.

(Robert Hooke)

EL-BADRY, M. A. (Institute of Statistics, Cairo Univ.) 4.1 (4.3)
The optimum estimate of moments—*In English*
Egyptian Statist. J. (1957) **1**, 33-38 (2 refs.)

The problem of estimating μ_k, the *kth* central moment ($k>1$) of a distribution, of which a set of statistically independent observations is known, is considered.

The author defines the optimum polynomial estimate as a polynomial in the observation such that its expectation is equal to μ_k and which has the minimum variance in the class of unbiased estimates.

He shows that such an estimate must be

1. a symmetric polynomial in the observations
2. homogeneous of degree k
3. and that only one of the observations in each of the terms of the polynomial can have a power greater than one.

For this he makes the assumption that the moment of degree $2k$ exists and gives an explicit expression for the estimate.

He applies his result to give simple expressions for the estimate of the variance, and the estimates of the third and fourth central moments.

(J. B. Simaika)

FERGUSON, T. (University of California, Los Angeles) **4.2** (2.4)

A method of generating best asymptotically normal estimates with application to the estimation of bacterial densities—*In English*

Ann. Math. Statist. (1958) **29**, 1046-1062 (18 refs.)

The minimum χ^2 method of estimation introduced originally by Karl Pearson and designed for use with the multinomial distribution has been extended and modified by Neyman and his co-workers in the past several years. These modifications have been introduced to simplify the calculation procedures while retaining the properties that these estimates have in common with maximum likelihood estimates, viz. asymptotic normality and minimum asymptotic variance (under suitable regularity conditions). Estimates with these properties are called best asymptotically normal estimates. A review of four modifications of the minimum χ^2 method is given both for the multinomial distribution and for a more general situation. References are given to papers giving examples of the application of these methods.

In all of these methods the estimation procedure involves the minimisation of a quadratic form in the observations. The author shows that it is also possible to start directly with a linear form and obtain best asymptotically normal estimates and, further, this procedure may be coupled with any of the usual modifications.

The procedure developed is applied to the problem of estimating bacterial density by the dilution method. In this method, from a large volume of the liquid containing λ bacteria per unit volume (λ unknown), samples of size α_i unit volumes are withdrawn and tested for presence or absence of bacteria. The probability that no bacterium appears in a single sample of volume α_i is $p_i = e^{-\alpha_i\lambda}$. The estimate of λ is based on the number of sterile samples and since p_i is an exponential function of λ, the maximum likelihood estimate is the root of a complicated non-linear equation. References are given to papers where methods to solve the equation are discussed. The author's modification of minimum χ^2 estimation still leads to an exponential equation but one claimed to be easier to solve than that yielding the maximum likelihood estimate. Numerical studies of the small sample properties of the new estimate are promised for a later paper.

(D. G. Chapman)

FINNEY, D. J. (University of Aberdeen & A.R.C. Unit of Statistics) **4.3** (10.7)

The efficiencies of alternative estimators for an asymptotic regression equation—*In English*

Biometrika (1958) **45**, 370-388 (9 refs., 4 tables)

This paper investigates the properties of various methods of estimating the parameter in the regression

$$E(y)=\alpha-\beta\rho^x.$$

Two models are considered for the generation of the deviations : (*a*) the y are independent and normally distributed about their expectations with constant variance, and (*b*) the y arise from a continuous autoregressive process in which the variance increases with x and successive values of y are correlated.

The method of statistical differentials is used to obtain general expressions for the bias and variance of a ratio. In the further investigation, detailed results are obtained only for the case (case 4, say) of four observations equally spaced in x. For this case, explicit expressions (in terms of ρ) are obtained for the following estimators, for both the models (*a*) and (*b*) above ; these are evaluated in four tables.

(i) Patterson's linear estimators. These are the ratios of weighted averages of the y (omitting the first) to the same weighted average of the y (omitting the last). In case 4, there is one degree of freedom in the weights, and three values (1, 1·25, 2) of an arbitrary parameter are suggested as being reasonable for different ranges of ρ. This estimation procedure has high asymptotic efficiency (as the variance of the y approaches zero).

(ii) The estimate obtained by performing the standard regression of y_{x+1} on y_x. For model (*b*), this gives the maximum likelihood estimator. In case 4, model (*a*), it has high efficiency, and the bias is of the same order as for (i), though of opposite sign.

(iii) Two estimates proposed by Hartley, using more complicated autoregressions of the y. In case 4, in comparison with (i) and (ii), these have similar variances but rather larger biases. [Biases not evaluated for model (*b*).]

(iv) A "general quadratic estimator" (ratio of two quadratic functions of the y) of which in case 4 a particular example is chosen for study. For model (*a*), this has high efficiency (bias not evaluated) but has no special merits.

This study is continued by Patterson, *Biometrika* (1958) **45**, 389-400 (q.v.).

(C. L. Mallows)

GUTTMAN, I. (University of Alberta and Princeton University) **4.2** (2.5)
A note on the series solution of a problem in estimation—*In English*
Biometrika (1958) **45**, 565-567 (4 refs.)

Suppose t is a complete and sufficient statistic, assuming only non-negative integer values, having a probability distribution of exponential type depending on a parameter θ which takes values in an interval containing the origin. Then there exists an (essentially unique) unbiased estimate with uniformly minimum variance of any chosen function of θ if, and only if, a certain simple function is analytic at $\theta = 0$, and has a power series expansion whose rth coefficient is zero whenever the probability that t equals r is zero. The estimate is obtained very simply in terms of these coefficients.

This result facilitates the construction of unbiased uniformly minimum variance estimates without recourse to the usual method involving the evaluation of the conditional expectation of an arbitrary unbiased estimate. It gives a result in cases where the method given in the 1946 paper by Girshick, Mosteller and Savage [*Ann. Math. Statist.* **17**, 13] fails. The method is applied to the cases of negative binomial and Poisson distributions.

(C. L. Mallows)

HÁJEK, J. (Math. Inst., Czechoslovak Academy of Sciences, Prague) **4.3** (8.1)
On the theory of ratio estimates—*In English*
Aplik. Mat. (1958) **3**, 384-398 (9 refs.)

Having selected n sampling units from a finite population by simple random sampling, we wish to estimate the aggregate value of a variable, say y, under the supposition that we know *a priori* (from sources other than the sample the aggregate values of a closely related variable, say x, for a system of subclasses of the population. The border-lines between subclasses may cut the sampling units, as in the case when sampling units are groups of persons and classification is with respect to age. We use a ratio estimate which is the sum of simple ratios for separate subclasses multiplied by respective x-aggregates. The variance and estimated variance of this ratio estimate are derived and the corresponding confidence interval is considered. The discussion covers the more general case when a third variable, say z, is used, which (1) is useless where x is known (e.g. x are fresh data and z are past data), but (2) is such that the aggregate values are at our disposal for a superposed system of subclasses ; (see Hansen *et al.*, *Sample Survey Methods and Theory*, Vol. I, pp. 413, 569 ; Vol. II, p. 226).

The analyses proceed as follows : First, the ratio estimate is approximated by a linear estimate obtained by Taylor expansion with respect to the random variables which are involved and measured about their expectations. Then a simple approximate variance and estimated variance are obtained in the usual way, with the only modification that certain sums are treated as single observations. Second, under the condition that certain regression relations, suggested by the linear approximation, hold it is proved that the ratio estimate is conditionally unbiased, and the conditional variance and estimated conditional variance are derived. Third, a supposition of normality is added and it is shown that the usual Student's-form confidence interval using the estimated conditional variance t_a corresponding to $n-1$ degrees of freedom and confidence level α covers the true value with a probability *greater* than $1-\alpha$. Also, in the absence of subclasses, Fieller's confidence interval may be used. But it appears that, using the same t_a, Fieller's interval is always longer. This is caused by the fact that the Fieller's method makes no use of any assumptions about the type of regression. In sample survey conditions, however, it seems plausible to expect that the regression passes approximately through the origin, or, when subclasses are present, that regression relations supposed in the paper are fulfilled to a reasonable degree of an approximation.

The remainder of the paper deals with some other methods of sampling and estimating and with modifications needed when stratification and subsampling are present.

(Author's summary.)

HORÁLEK, V. 4.3 (2.3)

A contribution to the study of the structure of materials—*In Czech* *Aplik. Mat.* (1958) **3**, 376-383 (4 refs., 1 fig.)

The author considers a specimen of metal in which small particles of a particular structural component are of interest. This body is cut through and on the cross-section there are observed the diameters of intersections of particles and the average number of intersections per unit area. It is desired to obtain an estimator for the average number of particles per unit volume. In the paper an estimator is found under the supposition that particles are distributed uniformly and independently and are of spherical shape with a log-normally distributed diameter. Then the conditional distribution of diameters of an intersection, under the condition that the particle has been intersected, is also log-normally distributed, and an estimator may be obtained which is a product of the sum of reciprocal values of diameters, of the average number of intersections per unit area and of a constant. The estimation procedure is adjusted for the case when only n randomly chosen diameters are measured and intersections are counted in only m randomly chosen unit areas. A ratio m/n which minimises costs is found. A numerical example and a photograph showing a typical cross-section are presented.

(J. Hájek)

LE CAM, L. (A. P. Sloan Foundation) 4.0 (1.8)

The asymptotic properties of Bayes solutions—*In French* *Publ. Inst. Statist. Paris* (1958) **2**, 17-35 (7 refs.)

This paper summarises and extends some work done in this field during 1953 by the same author. His object in this present paper is to give a precise statement to the effect that the choice of the *a priori* distribution tends to become weaker when the number of observations increases indefinitely. This statement can be found on page 21 of the paper and the whole investigation is set within the framework of Wald's decision theory.

(D. Dugué)

LEJEUNE, J. (Université de Paris) 4·3 (2.8)

On an *a priori* solution of Haldane's *a posteriori* method—*In French*
Biometrics (1958) **14**, 513-520 (4 refs., 3 tables)

The binomial distribution with the zero observations truncated frequently occurs in human genetics. When data on families in which a defective (taré) trait is transmitted are available, frequently the families comprising the data are such that at least one of the offspring shows that trait. This is especially the case when the trait is governed by a recessive gene. Under these conditions estimation and tests of the proportion, p, of the offspring carrying the trait tend to be numerically tedious.

The author rejects the procedure of Bernstein and Hogben for testing $p = 1/4, 1/2$ and is concerned to simplify the application of the method of maximum likelihood estimation of p as developed by Haldane. The estimation equation is equivalent to that obtained by the method of moments and is a linear function of $p/\{1-(1-p)^n\}$; no explicit solution is available.

This paper presents tables for facilitating the computation of the maximum likelihood estimate and its estimated variance. The first table tabulates the above function to four decimal places fo r $n = 2(1)15$ and $p = 0{\cdot}60(0{\cdot}05)0{\cdot}05$ and $0{\cdot}0275$, $0{\cdot}0225$.

The use of the tables is as follows: an arbitrary value, usually $0{\cdot}25$, $0{\cdot}50$, is chosen for p and the expected number of abnormal offspring is calculated according to the formula. This is repeated until the two neighbouring values p_1, p_2, which give expected numbers straddling the observed number, are discovered. At this stage linear interpolation is accurate to within 1 per cent.

Haldane has also given an expression for the inverse of the variance of the estimator. The author rearranges this into a linear function of a function of p, n which is tabulated for the above values of p, n. The inverse variance is calculated for p_1, p_2 and that for the interpolated value of p obtained by linear interpolation on these quantities. Although this underestimates the variance slightly, it is within 1 per cent. and so sufficiently accurate for practical purposes.

The use of both tables is illustrated by application to some data on ataxia. The terms *a priori*, *a posteriori* are used by the author to refer to probability and relative likelihood calculations respectively. Thus the title seems to arise from the previously noted equivalence of maximum likelihood and moment estimators.

(C. W. Clunies-Ross)

MENDENHALL, W. & HADER, R. J. (North Carolina State Coll.) 4·3 (2.5)

Estimation of parameters of mixed exponentially distributed failure-time distribution from censored life test data—*In English*
Biometrika (1958) **45**, 504-520 (17 refs., 6 tables, 3 figs.)

This paper is concerned with the estimation of population parameters for sub-populations where the measured characteristic of both the sub-populations and the whole population is distributed exponentially. The particular field of application is in connection with life test data derived from sample observations censored with respect to the time at which the test is terminated. After a brief introduction, a statement of the population model and the method of sampling, the authors deal with the estimation of the parameters according to whether or not the relative magnitude of the sub-population parameter is known. The method of maximum likelihood is used and estimating equations are given for the case of two sub-populations, i.e., the parameter for each together with the proportion that the size of one bears to the whole. These estimating equations involve an additional factor for which an estimating equation is also given together with an iterative process for estimation based upon a modification of the MLE obtained in 1956 by Deemer & Votaw [*Ann. Math. Statist.* **26**, 498-504] for samples drawn from a singly truncated exponential distribution. In the case where the relative magnitude of the sub-population parameters is known and the given procedure yields estimates which reverse the relationship, an adjusted estimation procedure is given: this condition is termed a "crossover" of the estimated sub-population parameters.

The small sample properties of the estimates were obtained by empirical methods for a number of parameter points, i.e., combinations of sample size and value of the sub-populations' parameters. The results are shown in a full table and it is concluded that the estimation is poorest, both with respect to bias and variance, for those parameter points at which a large number of crossovers occurred: a substantial improvement is achieved by using the adjusted procedure. The large sample variances were obtained by inverting the symmetric information matrix and a short table is given comparing the asymptotic variances with the estimated variances for five parameter points. The effect of the level at which the sample is censored is considered in relation to the efficiency of the estimates—here in connection with the problem of whether the cost of an increase in the testing time will be offset by enough additional information.

After showing that the results for the mixed exponentially distributed sub-populations are but a logical extension of those for a single exponential distribution, the authors generalise their model to a situation for which the exponential distribution is a particular case. The model chosen is the well-known family of distributions associated with the name of Weibull. The authors state that it should not be too difficult to set up an iterative method for the solution of the ML equations provided the number of sub-populations is small: reference is made to the use of standard numerical methods for the solution of equations by iteration with or without the use of a computer.

(Wm. R. Buckland)

PATTERSON, H. D. (Rothamsted Experimental Station, Harpenden, England) 4.3 (10.7)

The use of autoregression in fitting an exponential curve—*In English*

Biometrika (1958) **45**, 389-400 (5 refs., 4 tables)

This paper complements the investigations of Finney, *Biometrika* (1958) **45**, 370-388 (q.v.) into the properties of certain estimators of the "base" parameter ρ in an exponential regression. The residuals are assumed to be independent, normal with zero means and constant variances.

(i) The simplest estimator considered [denoted by (ii) in the summary of Finney's paper] is that obtained by straightforward analysis of the regression of y_{x+1} on y_x.

(ii) More generally, the regression of y_{x+1} on a weighted average ky_x+ly_{x+1} is considered.

(iii) Finally, a "general quadratic estimate" (ratio of two quadratic functions of the y) is considered. It is not obvious whether this is more or less general than Finney's estimator (iv).

Using matrix notation, general expressions for the leading terms in the bias and variance of (iii) are obtained. The variance is independent of k and l. It is shown how to choose the estimator so as to minimise the variance for any chosen value of ρ, and suggestions are made for computing the estimator.

Taking ρ equal to zero [which gives (ii) above], the efficiencies of these estimators are tabulated for the cases of 4, 5, 6, and 7 observations, and are compared with those of Patterson's "linear" estimators [(i) in the abstract of Finney's paper]. It is found that the "regression" estimator is substantially less efficient than the "linear" estimator in the case of 7 observations, over the range of the most useful values of ρ. The bias of the "regression" estimator is a function of k and l, is tabulated for case (i) above, and is compared with that of the "linear" estimator. The bias is large, and becomes more serious with increasing numbers of observations. It can be controlled by suitably choosing k and l, and is exhibited for some specimen values. The choice is difficult unless the number of observations is small.

The conclusion is that in general the regression method is unsuitable. Preliminary results assuming the residuals to be autocorrelated [model (*b*) in the abstract of Finney's paper] indicate that Steven's method (*Biometrics*, 1951, **1**, 247-267) can safely be recommended for most biological applications.

(C. L. Mallows)

RADHAKRISHNA RAO, C. (Indian Statistical Institute, Calcutta) 4.1 (4.3)

Maximum likelihood estimation for the multinomial distribution with infinite number of cells—*In English*

Sankhyā (1958) **20**, 211-218 (8 refs.)

In this paper the author extends his earlier results [*Sankhyā* **18**, 139] concerning maximum likelihood (M.L.) estimation to the case of a multinomial distribution with an infinite number of cells.

The author shows, under the sole restriction that $\Sigma\pi_i \log \pi_i > -\infty$ where the π_i represent the population frequencies in the different cells, that the maximum likelihood estimate of the population distribution converges to the actual distribution in the strongest of intrinsic senses. This result can also be proved by using the arguments of Wald [*Ann. Math. Statist.*, **20**, 595] as pointed out by Kiefer and Wolfowitz [*Ann. Math. Statist.*, **27**, 887] although the author's proof is simpler and more direct. In the parametric situation, the author proves, under the further restriction that the parameter to be estimated is a continuous functional of the distribution, that the approximate M.L. estimates corresponding to any fixed degree of approximation c $(0<c\leqq 1)$, are uniformly consistent.

In the next part of the paper, the author discusses maximum likelihood equation (M.L.E.) estimates, i.e. those which provide the maximum likelihood when the parameter is restricted to the roots of the likelihood equation. Under the additional assumption that there exists some neighbourhood of the true value of the parameter in which all the cell probabilities can be differentiated, the author demonstrates that, with probability one, for all sufficiently large n (n being the size of the sample), maximum likelihood estimates exist, that the class of M.L. estimates coincides with the class of maximum likelihood equation estimates, and that the M.L.E. estimates are uniformly consistent.

The author remarks that the regularity conditions imposed by him in the derivation of the above results seem to be considerably weaker than those assumed by earlier authors, and that the methods developed in the present paper apply without modification to the situation when the number of parameters to be estimated is greater than one.

(V. S. Varadarajan)

SARHAN, A. E. (High Institute of Public Health, Alexandria) 4.6 (2.5)
Estimation of central tendency for asymmetric distributions—*In English*
Egyptian Statist. J. (1957) 1, 7-14 (4 refs., 4 tables)

The author considers three measures of central tendency (location): the mean, the median and the mode in asymmetric distributions. He estimates these measures by the unbiased linear minimum variance method and compares the variances of the different estimates for three particular asymmetric distributions.

He shows that for 1. the two-parameter single exponential distribution and for 2. the one-parameter single exponential distribution, the variance of the sample median is approximately one-half the variance of the sample mean. While for 3. a skewed distribution with finite range (the mode dividing the range in the ratio 2 : 1), he shows that the variance of the sample mode is slightly smaller than the variance of the sample median and that this last variance is slightly smaller than the variance of the sample mean.

(J. B. Simaika)

SHENTON, L. R. (College of Science and Technology, Manchester) 4.1 (4.3)
Moment estimators and maximum likelihood—*In English*
Biometrika (1958) **45**, 411-420 (16 refs., 2 figs.)

A probability distribution function $p(x\,;\ \theta_1 \ldots \theta_h)$ is dependent on h unknown parameters, is continuous in x, possesses all moments and is such that the range of x is independent of $\theta \equiv (\theta_1, \ldots, \theta_h)$. The polynomials $\{q_s(x)\}$ dependent also on θ, are defined by their orthogonality with respect to expectations over $p(x\,;\ \theta)$. A series of these polynomials (to order r) is fitted (by moments) to the differential with respect to θ_j of $\log\ p(x\,;\ \theta)$. The fitting polynomial is denoted $A_j^{(r)}(x)$ and is sample mean value, denoted $\mathscr{A}_j^{(r)}$, is then a linear sum of the first r sample moments (whose coefficients are functions of θ) which has zero expectation. The equations $\mathscr{A}_j = 0, j = 1, \ldots\ldots, h$ are then solved for θ to get estimates $\hat{\theta}_1, \ldots \hat{\theta}_h$, which are functions of the first r sample moments. By what is in effect an application of Cramer's theorem on functions of sample moments, the leading terms in the expansion of cov. $(\hat{\theta}_i, \hat{\theta}_j)$ as power series in $1/N$ are obtained and these are shown formally to approach the corresponding maximum likelihood results of Fisher as $r \to \infty$. The approach to the limit is shown to be monotonic for $h = 1$ and the corresponding result for var. $(\hat{\theta}_i)\, i = 1, \ldots\ldots h$ is stated to hold generally (a method is given which is said to provide the outlines of a proof). Three distributions are taken to exemplify the method. The first is the truncated Poisson where the method yields David & Johnson's sufficient moment or maximum likelihood statistic. The second is an artificial example (the product of e^{-x} by a function linear in x and the unknown parameter) where the estimator turns out to be the root of an rth order polynomial. The third is the negative binomial of mean λ and variance $\lambda(1+\lambda/\alpha)$. The estimator of λ turns out to be the usual maximum likelihood estimator (the sample mean) and $\hat{\alpha}$ the root of a $(2r-3)$rd order polynomial. This, the author states, he is unable to show has a unique root but he shows that for r as small as 4 his estimator has high efficiency for α relatively small. The paper closes with a brief discussion of the analytic conditions on $p(x\,;\ \theta)$ necessary to ensure that the efficiency of the estimators tends to unity as $r \to \infty$ but no useful conclusion is reached. There is a final section referring to earlier papers of the author's where analogous problems arise. The author does not investigate the small sample behaviour of these estimators in this present paper.

(D. E. Barton)

TANNER, J. G. (Road Research Lab., D.S.I.R., England) 4.3 (5.2)
A problem in the combination of accident frequencies—*In English*
Biometrika (1958) **45**, 331-342 (2 refs., 3 tables, 2 figs.)

The paper is concerned with the accident frequencies in a period before and after similar changes in road conditions have been made at a number of sites. Comparisons are made using a set of control areas for which it is assumed that the ratio of accidents in the 'before' and 'after' periods is free of error. This use of control sites permits that the period of observation on the changed sites may differ. The total number of accidents occurring at a given changed site is then assumed to divide binomially into the number of accidents in the before and after change periods, the probability parameter of the binomial distribution being expressed as a function of κ_i which is a measure of the change at the ith site and takes the value unity when no change occurs.

In the first instance it is assumed that there is no variation between the values κ_i at N sites and maximum likelihood applied to estimate the common value κ and to test the hypothesis $\kappa = 1$, i.e. that no change in accident frequency has been observed.

Later an expression which measures the variation between the κ_i is constructed and which may be used to test the hypothesis that the κ_i have a common value, not necessarily unity.

So that inference may be made with regard to all sites which undergo the similar change in road condition rather than to the N sites which are under direct investigation, it is finally supposed that a parent population exists and that the values of κ_i form a sample from this population. Estimates of the first and second moments of the parent population are proposed together with a method for testing the hypothesis that the population first moment is unity.

In an appendix the first four moments of the logarithm of the estimated population first moment are given to first order when all the κ_i are equal.

(J. G. Saw)

WEILER, H. (Commonw. Sci. & Ind. Res. Organisation, Prospect, N.S.W.) 4.4 (3.6)
Confidence limits for the mean of a normal population with known coefficient of variation—*In English*
Aust. J. Appl. Sci. (1958) **9**, 321-325 (1 ref., 2 figs.)

Let $\bar{x}$ be the arithmetic mean of a random sample of size n from a normal population whose mean μ and variance σ^2 are unknown. If the coefficient of variation $k = \sigma/\mu$ of the distribution is known, the confidence limits for μ are determined by $\bar{x}/(1+A)$ and $\bar{x}/(1-A)$, where $A = Tk/\sqrt{n}$, and the α confidence interval for the standardised normal variate is $(-T, T)$. These limits depend only on k, $\bar{x}$, and n.

Let

(i) l_0 be the width $2Tk\mu/\sqrt{n}$ of the shortest confidence interval for μ, when the standard deviation σ is known ;

(ii) l_1 be the corresponding width $2ts/\sqrt{n}$ of the confidence interval when σ is unknown, t being the value corresponding to T obtained from Students' distribution for the sample of size n, and s the standard deviation of the sample ;

(iii) l be the width $2A\bar{x}/(1-A^2)$ of the confidence interval when k is known, using the above confidence limits. Then by calculating the expected values λ_0, λ_1 and λ of l_0, l_1 and l respectively, the efficiency of confidence limits based on l is compared with those based on l_0 and l_1.

It is found that λ_0 is always the shortest of the three. For small values of k, λ and λ_0 differ only slightly, while the difference between λ_0 and λ_1 is much larger. Only for k larger than 0·5 (approximately) does λ_1 turn out to be smaller than λ. For large values of n, both λ and λ_1 approach rapidly. By plotting $100(\lambda-\lambda_0)/\lambda_0$ and $100(\lambda_1-\lambda_0)/\lambda_0$ against different values of k, the expected values λ, λ_1, and λ_0 are compared graphically for samples of size 4, 6, 8, 10 and 30, and different values of k, for 95 and 99 per cent. confidence intervals.

An expression for the approximate expected value of s, the sample standard deviation is given in an Appendix, with an upper bound for the error involved.

(B. R. Bhat)

ADKE, S. R. & DHARMADHIKARI, S. W. (University of Poona) 5·7 (2.7)
Gain due to sequential sampling from a Gamma population—*In English*
Bull. Calcutta Statist. Ass. (1958) **8**, 91-94 (4 refs., 1 table)

It is well known that the use of sequential probability ratio test (SPR test) usually results in considerable saving in the average number of observations required to reach a decision, compared to that required by the classical Neyman-Pearson test of the same power. Wald has evaluated this gain, when observations are drawn from a Normal population with a known variance, to test the simple hypothesis H_0 (mean $\theta = \theta_0$) against the simple alternative H_1 ($\theta = \theta_1$). In this note the authors obtain the corresponding results for the Gamma population with mean $l\sigma$ and variance $l\sigma^2$ for testing the hypothesis H_0 ($\sigma = \sigma_0$) against H_1 ($\sigma = \sigma_1$) assuming that l is known. It is proved that the percentage gain in the number of observations is independent of l. Hence, if l is taken as unity, the problem can be treated as one pertaining to the exponential population with mean σ. This simplified problem is treated by Epstein and Sobel [*J. Amer. Statist. Assoc.*, **48**, 486, and *Ann. Math. Statist.*, **26**, 82]. The authors have reproduced the numerical results obtained by them for some values of $K = \sigma_0/\sigma_1$, α and β (the first and second kind of errors). The percentage gain, in these particular cases, is seen to range from 32 per cent. to 78 per cent.

(K. Madhava Sarma)

AKAO, Y. (Tokyo Institute of Technology) 5.2 (8.9)
A significance test for fraction-defective by compressed limits—*In Japanese*
Statist. Qual. Cont. (1958) **19**, 54-62 (8 refs., 1 table, 11 figs.)

When the probability that the variable under investigation exceeds the specification limits (in which case the item with this value is assumed to be defective) is very small, a large number of observations are needed to detect the change of the population mean of the variable by attribute; for example, by " go-not-go " device. It is well known that, in such a situation, compressed limits can be effectively used instead of the original specification limits in order to decrease the number of observations used in the control chart (cf. B.S.S. 1313 (1947): *Fraction-defective charts for Quality Control*).

The author has applied this device to the significance test of the null hypothesis that the population mean has a given value against the one-sided alternative based on the number of defectives and, given the method of obtaining the compressed limit which gives the minimum number of observations, to detect the given difference of two values of the population mean—under the null and alternative hypotheses for the given values of the probability of the first and second kinds of errors. The distribution of the basic variable is assumed to be normal with known variance.

The main point of this method is to determine the compressed limit so that it maximises the difference of two values of the probability that the variable exceeds the compressed limit under the null and alternative hypothesis. This leads to the simple result that the desired compressed limit is half the value of the difference of the two values of the population mean mentioned above, divided by the population standard deviation. The proof of the sample size being minimum, when the compressed limit is so determined, is not rigorous but is based on the normal approximation of the binomial distribution using the arcsine transformation. He also gives a graphical method of determining the number of observations and the critical number of defectives to test the hypothesis by using the compressed limit so obtained for fixed probabilities of the first and second kinds of errors.

(This paper was written in Japanese, but the translated paper in English will be published by the Union of Japanese Scientists and Engineers in the near future.)

(M. Siotani)

BEHRENS, W. U. (Kali-Chemie A.G., Hannover) 5.3 (6.4)
A note on discriminatory analysis—*In German*
Biom. Zeit. (1959) **1**, 3-14 (3 refs., 4 tables, 2 figs.)

This paper is concerned with the information supplied by inclusion of additional characters in problems of allocation and discrimination in the two-sample case. First the discriminatory power of one character is defined in terms of the standardised difference between sample means. Then it is shown that the additional discriminatory power supplied by a second character using a discriminant function is the greater, the higher is the absolute value of the coefficient of correlation between the two characters and the larger the difference between this correlation coefficient and the ratio of the second to the first standardised mean difference. Corresponding results follow for the addition of a third character. This statement, which is perplexing at first sight, is achieved by reformulating Mahalanobis' D^2 as a function of the standardised sample mean differences and correlation coefficients between the characters. It seems to be valid only under the assumption (not explicitly stated) that there is no inter-character correlation between sample mean differences. The derived formulae together with a table make it possible to judge the usefulness of additional characters prior to an evaluation of the discriminant function.

(R. Wette)

CLARINGBOLD, P. J. (University of Sydney, Australia) 5.8 (6.8)
Multivariate Quantal Analysis—*In English*
J. R. Statist. Soc. B (1958) **20**, 398-405 (13 refs., 3 tables)

In many biometrical investigations it is not possible to measure response on a continuous scale and recourse must be made to quantal scoring of signs exhibited by the animal. The author describes the canonical analysis of such multiple quantal variates. Considering q quantal variates $(y_1, y_2 \ldots y_q)$ the response vector is specified by 0 or 1 for each variate and the reaction pattern designated by the binary number it determines. Transformations of the original variables are considered which include the possibility of a reduction in the number of variates.

Suppose an experiment is performed so that the three sources of variation may be examined: (i) experimental error, (ii) parametric or treatment effects and (iii) residual effects. The dispersion matrices E and P are obtained by analysis of covariance of the variables y into the first two sources of error. Determination of canonical variates, that is, estimation of the required matrix of transformation, enables the use of variables with unit standard deviation and which are mutually uncorrelated. Successive rows of the transforming matrix obtained from the given equation correspond to the roots of the determinantal equation $|P-\Phi E|=0$. In the absence of treatment effect in a canonical variate, the corresponding root should approach unity and may be tested by an approximation to the F-test. Thus the minimum number of variates required to describe significant treatment variation is determined and the treatment variation is maximised in the derived system.

One special case of importance is considered in which a score is obtained which gives the maximum dependence of response on dose. A numerical example of an oestrogen assay is analysed in a 3^4 factorial experiment in which there are three quantal variates. Two roots of the determinantal equation are significant and corresponding estimated standardised vectors are obtained. In the experiment one of the factors was dose and the single variate which maximises dependence of response on dose is also obtained.

The author mentions that some workers have reduced a multivariate quantal problem to a univariate problem by pooling of classes and then applying standard methods of quantal analysis. He discusses the increase in precision resulting from the use of multivariate scores for response patterns, and stresses that, though the methods are somewhat laborious by manual computation, the availability of programmes for automatic analysis greatly encourage their adoption.

(N. W. Please)

COHEN, L. (City College of New York) 5.2 (4.4)

On mixed single sample experiments—*In English*

Ann. Math. Statist. (1958) **29**, 947-971 (4 refs., 3 tables, 9 figs.)

A mixed single sample (MSS) procedure is one which is derived from fixed sample (FS) procedures by choosing the sample size at random. For many problems use of MSS procedures will effect savings. For example, in testing two simple hypotheses the triple (α, β, n), where α (β) is the probability of Type I (II) error and n is the expected number of observations, is used as the operating characteristic (*OC*) of an MSS test (MSST) and in several problems considered by the author (including that of testing two means of a normal distribution with variance known) for any FST whose *OC* is (α, β, n) there is an MSST whose $(\alpha', \beta' n')$ is such that $\alpha' \leqq \alpha$, $\beta' \leqq \beta$, $n' \leqq n$ with strict inequality in at least one of these three. In the normal problem the good MSST's are randomisations of two FST's, the first taking O observations and accepting one of the hypotheses always, the second being a Neyman-Pearson test based on taking a non-trivial number n_o, of observations. In testing two rectangular distributions the first on $(\theta_1, 1+\theta_1)$, the second on $(\theta_2, 1+\theta_2)$ the author shows that the FST's are already the best one can do.

The author characterises the admissible (in the *OC* sense) MSST's for simple hypothesis testing by observing that an admissible MSST minimises for some positive a, b, c the quantity $a\alpha+b\beta+cn$ from which it follows that an admissible MSST is a randomisation over ν of FST's of sample size ν which are Bayes' solutions within the class of FST's of size ν and where the *a priori* distribution remains the same for all ν. A prescription for computing the MSST is then given. Some examples are given on how to obtain the best MSST for given α, β or for given (α, n) $((\beta, n))$.

The author remarks upon certain peculiar features of these MSS procedures when extended to confidence interval estimation and composite hypotheses testing namely, that an MSS confidence procedure will cometimes choose a confidence interval of O length and O confidence coefficient and that an MSST in the composite testing problem will not have a power function which tends to 1 as the parameter moves far away from H_o. This philosophical difficulty about choosing a sample of size O with positive probability also appears in the simple hypothesis testing problem.

(J. Sacks)

CUSIMANO, G. (School of Statistics, Palermo) 5.2 (2.5)

On conformity of statistical data to a probability distribution—*In Italian*

Ann. Fac. Econ. Com. Palermo (1958) **12**, 209-219 (1 ref., 2 tables)

In a statistical analysis the problem may arise of verifying the conformity of the observed data to the hypothesis that the empirical distribution of a particular phenomenon or character is according to a certain theoretical model. In this paper the author introduces some new statistics which aim at verifying the conformity of observed data to the hypothesis that the theoretical distribution is represented by the exponential Poisson law or by a new scheme of dependence recently constructed by Ferreri. The proposed statistics are :

$$\phi_3 = (x+2/x+1)(n_x n_{x+2}/n^2_{x+1})$$

and

$$\phi_4 = (x+3/x+1)(n_x n_{x+3}/n_{x+1} n_{x+2})$$

for the Poisson law where $n_r = (NM^r/r\,!)(e^{-M})$ and M is the arithmetic mean : N is the total frequency. For the Ferreri law we have :

$$\Phi_3 = [2-x(\phi_4-1)/1-x(\phi_3-1)]\,[\phi_3-1/\phi_4-1].$$

These statistics use only some groups of the observations, rather than the complete distribution, and therefore they can be used when the empirical distribution is truncated or censored. The statistics have been applied to examine the degree of correspondence to the Poisson of a distribution of 400 squares of an " emocitometer " arranged according to the number of observed cells and the conformity to the Ferreri distribution of the frequency distribution of 23,248 houses of Lione arranged according to the number of cancer deaths which occurred in each one during the period 1918-1938.

In the paper are also indicated some simple criteria of estimation for the parameter M of the Poisson law and of the parameters λ and d of Ferreri's model. These appear particularly interesting in the case of truncated or censored distributions as they, too, are based only on some of the groups of observations.

(F. Galantino)

DOORNBOS, R. & PRINS, H. J. (Unilever N.V. Rotterdam & N. V. Philips' Gloeilampenfabrieken, Eindhoven) **5.5** (5.6)

On slippage tests. III : Two distribution free slippage tests—*In English*
Proc. Kon. Ned. Akad., Wetensch. (1958) **61**
Indag. Math. (1958) **20**, 438-447 (6 refs., 2 tables)

This paper is the last part of a series of four papers on tests for slippage of one of a group of variates, or one of a group of means of sets of variates, the first three of which were published in *Indagationes Mathematicae* (1956), 329-337 and (1958), 38-55. In these papers a number of general methods are applied to the cases of slippage tests for variates following various specified distributions with unknown parameter values and to some special cases. The following methods were applied :

(*a*) A simple and close approximation method to the true level of significance.

(*b*) A method to obtain power functions in an approximate way [derived from method (*a*)].

(*c*) A test for slippage of one of a set of means of groups of independent observations of unequal size.

Method (*a*) was already introduced by Cochran. The present authors did not succeed in finding a general theory of slippage tests for one variate or one group mean. For this reason slippage tests for various cases are discussed in the papers. In the first papers, slippage tests for Γ-variates (or variance estimates from normal samples), normal variates, Poisson, binomial and negative binomial varieties were discussed.

The fourth and last paper gives first slippage tests for the largest and for the smallest of a number of column totals obtained by the method of m-rankings. It contains also a table for the cases $m=3(1)9$ and the number of ranked objects $k=2(1)10$.

The second test is a distribution-free k-sample slippage test for all sample sizes. The test statistic is obtained in the following way : all observations are ranked and each sample is tested with Wilcoxon's two-sample test (one-sided) against the pooled other samples. The sample, which is associated with the smallest probability of exceedance of the values of Wilcoxon's test statistic obtained, is said to have slipped if the probability concerned is smaller than the level of significance, chosen in advance, divided by the number of samples k.

In a way similar to that used by D. van Dantzig [*Indagationes Math.* (1951) **13**, 1-8] to prove the consistency of Wilcoxon's test against a certain class of alternative hypotheses, the consistency of this test procedure against a related class of alternatives is shown.

Finally, the paper contains a table of critical values for the test for Poisson variates described in the third part of this series of papers.

(The authors)

GNEDENKO, B. V. (Math. Inst., Academy of Sciences, Kiev) **5.3** (5.6)

On the Wilcoxon Test of comparing two samples—*In Russian*
Bull. Acad. Sci.—Sér. Sci. Math. Astr. et Phys. (1958) **6**, 611-614

A statistical test is called consistent if it distinguishes two different probability distributions with the probability tending to 1, when the number of experiments increases to infinity. It is easy to verify that the Kolmogoroff-Smirnoff test is consistent.

The paper contains an example proving that the Wilcoxon test is not consistent.

(L. Kubik)

Remark of the Regional Editor :

From the following additional references it is known that the Wilcoxon test distinguishes between the distribution functions $F(x)$ and $G(x)$ if and only if $\int_{-\infty}^{\infty} F(x)dG(x) \neq \frac{1}{2}$.

(*a*) D. van Dantzig (1951), On the consistency and the power of Wilcoxon's two-sample test (*Indagationes Math.*, **13**, 1-8).

(*b*) E. L. Lehmann (1951), Consistency and Unbiasedness of certain non-parametric tests (*Ann. Math. Statist.*, **22**, 165-180).

(*c*) A. Rényi (1953), Neue Kriterien zum Vergleich zweier Stichproben (*A Magyar Tudományos Akadémia Alk. Matematikai Int. Közleményei*, **2**, 243-265).

(A. Rényi)

ISHII, G. (Ikuno Senior High School, Osaka)
Kolmorgoroff-Smirnoff test in life test—*In English*
Ann. Inst. Statist. Math., Tokyo (1958) **10**, 37-46 (10 refs.)

5.6 (5.2)

The author considers certain distributions in connection with the Kolmorgoroff-Smirnoff distribution-free tests of goodness of fit. Let there be two independent random variables with empiric distribution functions $F_n(x)$ and $G_{nk}(x)$ with a common continuous distribution function $F(x)$. If the two sequences of observations from the empiric distribution functions are merged and ordered we have a sequence: $z_1 < z_2 < \ldots < z_m < \ldots < z_{n(k-1)}$. If a portion of this sequence is taken—as might occur if a life test was stopped at the mth failure—a theorem is proved on

$$P\{G_{nk}(x) - F_n(x)\} < \epsilon,$$

for $-\infty < x < z_m$. The method of proof is analagous to that used by Blackman in 1956 [*Ann. Math. Statist.*, **27**, 513-520]. The corresponding theorem relating to the opposite inequality, i.e., $> -\epsilon'$ (Theorem 2) is followed by two theorems (3 and 4) dealing with $P\{F(x) - F_n(x)\} < \epsilon$ for $-\infty < x < x_p$ and $P\{F(x) - F_n(x)\} > -\epsilon'$ for the same restricted range of the variate x. These four theorems are concerned with a one-sided alternative to a test of a null hypothesis relating to the distributions. Theorems 3 and 4 have an application to the case where a life test is terminated at a prearranged time x_p.

The test of the null hypothesis for the two empiric distributions (Theorem 5) contemplates a two-sided alternative and again the variable x is restricted to $-\infty < x < z_m$. In this paper only the case of $k = 1$ is treated, viz.:

$$P\{-\epsilon' < G_n(x) - F_n(x) < \epsilon\}.$$

Two corollaries are given to Theorem 4 and another for Theorem 5. For all the theorems and the corollaries the expressions for the probability distributions considered are very complex.

(Y. Suzaki)

KEMP, K. W. (British Nylon Spinners Ltd., Pontypool)
Formulae for calculating the operating characteristic and the average sample number of some sequential tests—*In English*
J. R. Statist. Soc. B (1958) **20**, 379-386 (3 refs., 4 tables, 4 figs.)

5.7 (4.7)

Work carried out by Page in 1954 (*J. R. Statist. Soc.* B. **16**, 36-139) in connection with a sequential sampling scheme enables approximations to be found to the operating characteristic and the average sample number. The present author takes up this work, uses the same method as Page but makes different assumptions, and obtains closer approximations. The method is applicable to a broad class of distribution functions for the random variable characterising the unit of the sample, but calculations are performed for the normal case only. Comparisons are made between the proposed method of approximation, Page's method and the true values.

(F. N. David)

LEHMANN, E. L. (University of California, Berkeley) **5.1 (5.2)**

Significance level and power—*In English*

Ann. Math. Statist. (1958) **29**, 1167-1176 (13 refs., 2 figs.)

The problems considered are (*a*) the general problem of choosing, for a significance test based on a given experiment, among possible significance levels α in the light of corresponding possible test powers β; and (*b*) the corresponding problem in the presence of a nuisance parameter ν, particularly in the case of multiparameter exponential distributions. As a basis for treating the latter case, a minimal essentially complete class of tests is characterised for the problem of a one-sided test of a hypothesis $H_0 : \theta \leqslant \theta_0$ (against $\theta > \theta_0$) on one of the parameters. Denoting by t the (complete) sufficient statistic for ν when $\theta = \theta_0$, it is proposed to use conditional tests having conditional level $\alpha(t)$ at θ_0 and conditional power $\beta(t)$ at θ_1, where θ_1 is a chosen value of θ for which it is important to reject H_0.

Then the general problem (*a*) can be considered conditionally on an observed value of t. Several possible approaches to the general problem (*a*) are described and suggested as more reasonable in many cases than use of a conventional level such as ·05 or ·01 for α or $\alpha(t)$. One of these is the choice of some constant k not very far from unity, and the choice of a test which satisfies $\alpha = k(1-\beta)$, or of a conditional test which satisfies $\alpha(t) = k(1-\beta(t))$. Several examples are treated in detail.

(A. Birnbaum)

71

MORICE, E. (Inst. Nat. Statist. et Etudes Econ., Paris) **5.6 (5.3)**

Some non-parametric tests—*In French*

J. Soc. Statist. Paris (1958) **99**, 254-284 (16 refs., 11 tables, 1 fig.)

This paper summarises the work of recent years upon non-parametric two-sample tests which has been published in statistical (English-language) journals. It is given in three parts: (1) Tests based upon the sum or average of ranks, (2) tests based upon the number and length of runs of alternatives, (3) tests based upon the signs of successive differences.

In the first part the tests of Wilcoxon [*Biometrics*, 1945] Colin White [*Biometrics*, 1952] and Mann & Whitney [*Ann. Math. Statist.*, 1947] are reviewed together with the k-sample test of Kruskal & Wallis [*J. Amer. Statist. Assoc.*, 1952]. Examples are given to illustrate these tests and the use of the special tables, extracts of which are included in the text. The second part summarises the tests proposed by Mood [*Ann. Math. Statist.*, 1940], Mosteller [*Ann. Math. Statist.*, 1941], Sved & Eisenhart [*Ann. Math. Statist.*, 1943] and Wald & Wolfowitz [*Ann. Math. Statist.*, 1940 and 1944]. Extracts of tables from the paper by Sved & Eisenhart are given and reference is made to the tables by Olmstead [*Ann. Math. Statist.*, 1947].

In the third part, the author considers the tests proposed by Moore & Wallis [*J. Amer. Statist. Assoc.*, 1943] and Mann [*Ann. Math. Statist.*, 1945]. In this connection the paper reviews not only tests concerning the number of positive and negative signs (of differences) but also the length of runs of differences of the same sign: the tables by Olmstead are relevant in this latter instance.

(D. Dugué)

SALIB, S. R. (Faculty of Engineering, Ain Shams University) 5.6 (5.1)

Asymptotic power functions of some statistical rank order tests—*In English*
Egyptian Statist. J. (1957) 1, 15-32 (5 refs.)

Let X, Y, Z be three random variables and let $F(t)$, $G(t)$, and $H(t)$ their distribution functions be such that they differ only in the value of the location parameter. The problem of testing the null hypothesis $F(t) = G(t) = H(t)$ against the alternatives $F(t)<G(t)<H(t)$ for all t is considered.

To test this hypothesis, the author considers three samples (x_i) $i = 1, 2, \ldots, n_1$, (y_j) $j = 1, 2, \ldots, n_2$ and (z_k) $k = 1, 2, \ldots, n_3$ and introduces the two statistics :

1. The triplet statistic $U_{n_1n_2n_3}$ defined by the relation $n_1n_2n_3\ U_{n_1n_2n_3}$ equals the number of possible triplets (x_i, y_j, z_k) such that $x_i<y_j<z_k$.

2. The rank ordered statistic $L_{n_1n_2n_3}$ defined in the following way : Let (W_t) $t = 1, 2, \ldots, n_1+n_2+n_3$ be the ordered combined sample ; the statistic $L_{n_1n_2n_3}$ is the least number of nearest neighbour interchanges which turn (W_t) into a sequence of three runs, with the first run formed by the elements drawn from the first population, the second run from the second population, and the third run from the third population.

The author gives here the expectation and variance of the $U_{n_1n_2n_3}$ under the null hypothesis and shows that its distribution is asymptotically normal. And, as he had previously obtained the expectation and the variance of $L_{n_1n_2n_3}$ and had shown that its distribution is also asymptotically normal, he compares the asymptotic efficiency of the two tests.

He then shows that, in the particular case where the distribution of the variables is rectangular, the relative efficiency of L_{nnn} with respect to U_{nnn} is $2/15$. And, in the case where the distribution of the variables is normal, he shows that the relative efficiency is $0{\cdot}3$.

Lastly, the author notes that the number of comparisons used in the triplet statistic U has the order $O(n^3)$ while the number of pairs of comparisons in the L statistic is of the order $O(n^2)$.

(J. B. Simaika)

SUKHATME, B. V. (Indian Council of Agric. Res., New Delhi) 5.6 (1.3)

A two-sample distribution free test for comparing variances—*In English*
Biometrika (1958) **45**, 544-549 (4 refs.)

A two-sample distribution-free test is proposed for comparing variances of two populations and a general formula is derived for its asymptotic efficiency with respect to the variance ratio (F-test). The proposed test presupposes knowledge about the relative locations of the two populations. Where this information is not available it is suggested that deviations of the observations from the sample medians be considered and a modification of the test crition is proposed for this situation. It is shown that under certain regularity conditions the modified test is asymptotically distribution-free.

(F. N. David)

VINCZE, I. (Math. Inst., Hungarian Academy of Science, Budapest) 5.6 (5.3)

Some two-dimensional distributions and limit-distributions in the theory of ordered sampling—*In German*

Mag. Tud. Akad. Kut. Mat. Intézet. Közleményei (1957) 183-209 (15 refs.)

In this paper the author determines the joint distributions and limiting distributions of some pairs of statistics.

Let $\xi_1, \xi_2, \ldots, \xi_n$ and $\eta_1, \eta_2, \ldots, \eta_n$ be samples concerning the random variables ξ and η with common distribution function $F(x)$. Let us denote further with $F_n(x)$ and $G_n(x)$ the empirical distribution functions of these samples.

Let $\xi_0^{(n)}$ and $\eta_0^{(n)}$ denote in the following those smallest values of x, where the functions $F_n(x+0)-G_n(x+0)$ and $|F_n(x+0)-G_n(x+0)|$, respectively take their maxima.

In theorem 1, the author determines with the aid of two combinatorical lemmas the joint distribution of the following two statistics :

$$\sup_{-\infty<x<\infty} [F_n(x)-G_n(x)],\ \tfrac{1}{2}[F_n(\xi_0^{(n)}+0)+G_n(\xi_0^{(n)}+0)]$$

In theorem 2, the author considers the joint distribution of the statistics

$$\sup_{-\infty<x<\infty} |F_n(x)-G_n(x)|,\ \tfrac{1}{2}[F_n(\eta_0^{(n)}+0)+G_n(\eta_0^{(n)}+0)].$$

This problem is reduced to a random walk problem.

In theorems 3 and 4 the author determines the limiting distributions of the above distributions, in case $n\to\infty$. The marginal distributions of these limiting distributions are the well-known Smirnoff and Kolmogoroff-Smirnoff distributions, respectively relating to the one-sided and to the absolute deviation.

It is also shown that the limiting distribution of the variable $\xi_0^{(n)}$ agrees with the common distribution of the variables ξ and η.

Finally in theorem 5 the limiting distribution of the quotient of the variables

$$\sup_{-\infty<x<\infty} [F_n(x)-G_n(x)]$$

and

$$\tfrac{1}{2}[F_n(\xi_0^{(n)}+0)+G_n(\eta_0^{(n)}+0)]$$

is determined.

The author remarks that on the basis of these theorems, the tests for goodness of fit may be improved by using two statistics.

(E. Csaiki)

WALTER, E. (Max-Planck-Institut für Tierzucht und Tierernährung, Göttingen) 5.6 (5.3)

Some simple unbiassed non-parametric tests for the two-sample problem with pairs of observations—*In German*

Metrika (1958) 1, 81-88 (9 refs.)

A sample of size n consists of n pairs of numbers, each pair referring for example to the reaction of one person under two different conditions (treatments). For each pair, the difference of the two numbers is calculated. It is assumed that the differences are independent and that their distributions are identical and continuous. The null-hypothesis is that these distributions are symmetric with regard to the value zero, corresponding to the assumption that both conditions (treatments) are equivalent. The alternative consists of all continuous distributions which are *not* symmetric with regard to the value zero.

The test given below is uniformly unbiassed in the strict sense : the n differences are ordered according to their absolute magnitude and then divided into r groups ($r\geqq 2$), each group consisting of adjacent differences. Between every two adjacent groups one difference is deleted. The deleted differences are neglected in the whole procedure : n_1 denotes the number of differences in the ith group, d_1 the difference between the number of positive differences in the ith group and the number of negative differences in the ith group. Each test with the critical region $\Sigma g_i|d_i|^p\geqq H_0$; $g_i\geqq 0$, $p\geqq 1$ is uniformly unbiased in the strict sense. The proof of this theorem is given in a paper by E. Walter (so far unpublished).

Subsequently, some special cases are considered : $r=2$, $n_1=0$, $n_2=n-1$, $g_1=g_2=p=1$ is a test similar to the sign-test, the only distinction consisting in the fact that the difference with the smallest absolute amount is neglected. $r=2$, $n_1=n-k-1$, $n_2=k$, $g_1=0$, $g_2=p=1$ is a so-called maximum-test : the null-hypothesis is rejected, if the k differences with the greatest absolute amount are all of the same sign. In this case, the size of the test is 2^{-k+1}, independent of the sample-size.

Some special features of uniformly unbiassed nonparametric tests are discussed. Special consideration is given to the maximum-test whose power is compared with the power of the Wilcoxon (Mann-Whitney)-test and the sign-test in the special case of differences resulting from normally distributed variables with different means and identical variances.

(F. Ferschl)

BARTON, D. E. & CASLEY, D. J. (University College, London) **6.1 (4.3)**

A quick estimate of the regression coefficient—*In English*

Biometrika (1958) **45**, 431-435 (12 refs., 1 fig.)

In a bivariate sample $(x_i y_i)\, i = 1, 2, \ldots, n$ from a population in which the regression of y on x is linear the difference of the means of the k largest and the k smallest of the ordered x's is divided into the difference of the means of the corresponding y's to obtain an unbiased estimate of the slope of the linear regression. The efficiency of this estimate relative to the sample regression coefficient is found particularly for the normal bivariate distribution. In this latter case the estimate is shown to be the most efficient in large samples near the value $k = n \times 0{\cdot}27$ when the efficiency is 81 per cent. For small samples the first terms in an asymptotic expansion indicate an efficiency of between 70 per cent. and 80 per cent. for k between one-third and one-quarter of the sample size.

The distribution of the estimate is shown to be symmetrical and to tend to normality; the leading terms of an asymptotic expansion for the second moment ratio, β_2, of this distribution are obtained. The method used is to find the movements of the estimate, first conditional on x_1, $x_2, \ldots, x_n$ and then over $x_1, x_2, \ldots, x_n$ conditional on the upper and lower kth x-values. The expectation of these functions is obtained by the methods of David and Johnsón, "Statistical Treatment of Censored Data," *Biometrika* (1954).

(D. E. Barton)

BEJAR, J. (Instituto de Investigaciones Estadísticas, Madrid) **6.9 (5.2)**

Contingency tables—*In Spanish*

Trab. Estadist. (1958) **9**, 85-101

If quantitative observations of an aleatory variable are classified according to N distinct characteristics (quantitative or qualitative), the corresponding table of frequencies, which will be a double-entry table, is called a contingency table.

This work deals with problems related to the hypothesis of independence of all the characteristics or groups of characteristics which form the contingency tables. It aims to give a general method in order to solve the independence problems of any number N of classifications or characteristics or p sub-groups of characteristics. The case of two characteristics is first considered and a general sketch is given of the method of testing the hypothesis of independence of both classifications; according to the "Mathematical Methods of Statistics" of H. Cramér. This demonstration is generalised to the case of N classifications. The author then proceeds to consider the case of three characteristics. He puts the problem of testing the hypothesis that three of them should be independent according to the information given by a sample of n observations. The result is a generalisation of the former. It shall be coincident with it if one characteristic is taken without subdivision into intervals. The test ends with a distribution asymptotically convergent to the χ^2 distribution of Pearson with $rst+r-s-t+2$ degrees of freedom; where r, s, and t are the number of intervals into which each characteristic is respectively subdivided.

The next part of this paper considers the hypothesis of the independence of one characteristic (r intervals) as opposite to the other two (s and t intervals). The test is performed by means of a χ^2 test with $rst-r-st+1$ degrees of freedom: it coincides with the case of two characteristics if either $t = 1$ or $s = 1$. This case of the independence of a characteristic as opposed to the other two jointly considered is solved (as stated by the author) in *Introduction to the Theory of Statistics*. The test in which every characteristic is independent from the other is considered for the general case of N characteristics. The variation of each is resolved into $l_1, l_2, \ldots l_n$ sets or groups. The test is completed with a χ^2 test with appropriate degrees of freedom.

Finally, suppose that a group is arranged with the first h_1 characteristics, another with the second h_2 characteristics and so on until p groups are independent. The test leads asymptotically to a χ^2 distribution with $l_1, l_2, l_3, \ldots l_{h_1}+l_{h_1+1}, l_{h_1+2}, \ldots l_{h_2}+\ldots+l_{h_{p-1}+1}, l_{h_{p-1}+2} \ldots$ degrees of freedom.

(S. Rios)

COX, D. R. (Birkbeck College, London) **6.8 (5.2)**

Two further applications of a model for binary regression—*In English* *Biometrika* (1958) **45**, 562-565 (7 refs.)

Two distinct problems are considered, both dealing with the analysis of quantal data, i.e. data in which each observation takes one of two possible forms, 0 and 1, say. (Thus 0 may be " non-defective ", 1 " defective ".)

The first problem is concerned with matched pairs. Let there be n pairs of observations, the two observations in each pair corresponding say to two treatments A and B, there being in general a correlation within pairs in the response. For each pair let the odds of the response 1 for B be a constant, ψ, times the odds for A. The optimum test of the hypothesis $\psi = 1$ (i.e. that A and B are equivalent) is shown to be one due to McNemar. Confidence intervals for ψ are obtained ; if two or more independent sets of data are available, the equality of their ψ values is easily tested.

In the second problem, there are n independent trials and a sequence $p_1, \ldots, p_n$ is given. It is required to test whether the observations are consistent with the hypothesis that p_i is the probability of response 1 on the ith trial, $i = 1, \ldots, n$. Two statistics are introduced, one for examining whether the p_i are correctly spread, and the other for examining whether they are correctly centred. The two statistics can be combined to give a single χ^2 test with two degrees of freedom.

(D. R. Cox)

79

COX, D. R. (Birkbeck College, London) **6.8 (5.2)**

The regression analysis of binary sequences—*In English* *J. R. Statist. Soc.* B (1958), **20**, 215-242 (discussion, 34 refs., 1 table)

A sequence of 0's and 1's is observed and it is suspected that the chance that a particular trial is a 1 depends on the value of one or more independent variables. Tests and estimates for such situations are considered. The method is to assume a logistic dependence between the probability and the independent variable and, by considering an appropriate conditional distribution, to obtain an optimum exact test and estimate for the slope parameter of the logistic relation. The problems considered differ from those of, for example, bioassay, in that it is assumed that the slope of the logistic line is the parameter of interest, the position of the line being a nuisance parameter.

The following special cases are examined and related to earlier work :

(*a*) 2×2 contingency table ;

(*b*) $2 \times k$ contingency table with a quantitative variable associated with columns ;

(*c*) test for serial order effect ;

(*d*) test that the probability of occurrence is a specified value, against a trend-like alternative ;

(*e*) test that the probability depends on the result of the previous trial ;

(*f*) test that the probability depends on the total number of 1's in all previous trials.

In the discussion, examples of applications calling for these and more difficult methods are given, together with general and particular theoretical comments on the methods used.

(D. R. Cox)

COX, D. R. (Report of Discussion on paper by) **6.8** (5.2)
The regression analysis of binary sequences—*In English*
J. R. Statist. Soc. B (1958) **20**, 232-242.

The discussion on this paper consists of eight verbal contributions, one of which was extended in writing, together with an additional contribution in writing received after the meeting and the author's detailed reply. Those who took part, in order of speaking, were :—C. A. B. Smith, G. H. Jowett, Miss V. R. Cane, D. V. Lindley (subsequently extended in writing), D. E. Barton, H. E. Daniels, G. M. Jenkins and G. A. Barnard. The written contribution was from A. Stuart.

After commenting upon the possible applications of the paper, Dr Smith suggested that the difficulties of using the Gaussian or similar approximation to positively restricted variables might be overcome by using the distribution of the square of Gaussian variable with non-zero mean.

Dr Jowett and Miss Cane also commented upon application of the methods of the paper and Mr Lindley continued by offering a justification for the use of a conditional distribution in the argument of Section 4.1 of the paper. The justification rests upon the point that a desirable conditional test is a generally desirable test.

Dr Barton drew attention to the surprisingly close parallel between the present paper and results in the theory of " pure ranking " as developments of paired comparisons. He also drew attention to earlier work on the Wald-Wolfowitz runs test in connection with the binary Markoff chain by F. N. David [*Biometrika* (1947) **34**, 335-339] as developed by Barton and David [*Biometrika* (1957) **44**, 168-178] and Barton, David and Mallows [*Biometrika* (1958) **45**, 166-177].

Professor Daniels showed that even in the case of random sampling from a finite population the logistic law is a natural approach, apart from its simplicity and mathematical convenience, if the required approximate distribution is obtained by the method of steepest descents. Dr G. M. Jenkins contributed a simpler solution to the problem discussed in Section 6 of the paper—a solution which avoided any arbitrary assumption such as that of the logistic law. A criterion is presented which is similar to the non-circular serial covariance of lag one with a mean correction : raising the possibility that binary sequences, as considered in this paper, may be handled by a suitable modification of the criterion used in the autocorrelation analysis of continuous variables.

Professor Barnard pointed out some cautionary points to be observed in using the logistic—after remarking upon the value and simplicity of the logit transformation. Mr Stuart's written contribution, *inter alia*, draws attention to some additional points regarding the optimum character of the Wilcoxon Test for location problems.

The author's detailed reply offered further comments on points raised by the discussion.

(Wm. R. Buckland)

DANIELS, H. E. & KENDALL, M. G. (Univ. of Birmingham and London School of Economics) **6.2** (2.2)
Short proof of Dr. Harley's theorem on the correlation coefficient—*In English*
Biometrika (1958) **45**, 571-572 (4 refs.)

It has been shown by B. I. Harley that the mean value in repeated sampling of the inverse sine of the sample correlation coefficient is equal to the inverse sine of the population correlation coefficient. This result is proved by Daniels and Kendall using a combination of geometry and concordance. Their method of proof is much shorter than that given by Harley and demonstrates effectively the basic reason for the existence of the relationship.

(F. N. David)

HODGES, J. L., Jr. (University of California, Berkeley) **6.7** (0.2)

Fitting the logistic by maximum likelihood—*In English*
Biometrics (1958) **14**, 453-461 (6 refs., 3 tables)

Fitting a logistic response curve to quantal data by maximum likelihood usually entails an iterative process. The " transfer method " is presented as a means for eliminating or shortening the iterative process, depending on the degree of accuracy required. Common practice in logistic curve fitting involves sight fitting a straight line to points plotted on a logistic scale or on logit paper. Estimates of the parameters of the logistic curve from such a line are subjective and not well defined, serving only as the first step in the iterative process. The transfer method is based on the maximum likelihood estimates of the minimal sufficient statistics for the parameters, namely, the total response, estimated by the sum of the observed (positive) responses in several groups subject to different stimuli levels, and the total " moment ", estimated by the sum of the products of the stimuli and their corresponding observed responses.

The method is graphical and requires a straight line sight fitted to the original data as before. The aim is to determine a new set of responses, corresponding to the stimuli, which are collinear and which preserve the observed total response and total moment. This is accomplished by a series of transfers of responses from points above the line to points below the line, subject to the foregoing conditions. A new sight line is drawn and the process repeated until the points are collinear to graphical accuracy. Computers, tables or charts are not required. If graphical accuracy is sufficient, the parameters of the logistic curve may be determined from this final line. Otherwise these values may be used to initiate a single cycle of iteration which is claimed to result in accuracy usually attained only after several cycles. The transfer method assumes that the observations follow the logistic binomial model. The fitted line may be used to carry out a chi-square test (d.f. = no. levels − 2) of this assumption using the original unadjusted data. The transfer method is easiest to apply in cases where an equal number of trials is made at each level of stimulus and where the levels of stimulus are equally spaced. Examples are given for this case and for one in which the number of trials is not the same for each level of stimulus. The method is still applicable, though less simple, when successive levels of stimulus are not equally spaced. A third example is given to show the ease with which this method handles the zero or 100 per cent. response.

This paper deals only with the technique of maximum likelihood estimation and not with its desirability for fitting the logistic. Where the maximum likelihood estimates do not exist for certain values of response, it is also impossible to find transfers which preserve the total moment. The transfer method requires the existence of simple sufficient statistics and thus is not directly applicable to the fitting of the integrated normal. Nevertheless it provides a good set of starting values for the iterative process in this case. It is noted that probit instead of logit paper may be used with somewhat less accuracy and that methods such as the Spearman-Kärber technique may be used in lieu of the initial sight fitted line. Reference is made to extant and forthcoming charts and tables by Berkson which in part make this an interim method. (M. F. Peck)

KONIJN, H. S. (University of Sydney) **6.6** (10.0)

A restatement of the conditions for identifiability in complete systems of linear difference equations—*In English*
Metroeconomica (1958) **10**, 182-190 (3 references)

In Cowles Monograph No. 10, Koopmans, Rubin and Leipnik stated and proved certain theorems relating to identifiability in complete systems of linear difference equations where those equations were subject to disturbances, i.e. had stochastic error terms. The author, in using these theorems, has found that not only may the concept of identification as stated not be precisely the one required but that the theorems tend to lead to a certain amount of confusion. He suggests that some attention might usefully be given to conditions which make a set of coefficients either " just identifiable " or " unidentifiable ".

After giving particulars of his notation the author states his problem as being concerned with methods of testing whether the specification of any given model is sufficiently narrow to allow identification of the particular (but not fully known) structure or the identification of a certain set of its characteristics. The general sharpening of the concept of identification used in this paper rests on the definition of structures which are observationally equivalent. Two structures are " observationally equivalent " if, for the observed variables of the set of initial variables and for the reigning joint distribution of exogenous variables, the structures give rise to the same (conditional) distribution of the observable variables.

The author then states that the linearity of the model leads to the expectation that observational equivalence is closely related to invariance under non-singular linear transformations. After stating a lemma dealing with a necessary and sufficient condition, the author gives his own theorem—implied by this lemma and the definition of identifiability. He then states that the simplest application is the case in which the model specifies B_0 in $Ax_t = B_0 y_t + \Gamma(z_t) = u_t$ to be the identity matrix, I, and proceeds to apply his theorem to recursive models. Further sections of the paper investigate the identifiability of rows of the matrix A and the effect of homogeneous restrictions to single rows.

(Wm. R. Buckland)

LEVIS, V. (Faculty of Economics, University Institute, Venice) **6.1 (–.–)**
Mean Square Regression linear spaces—*In Italian*
R. C. Mat. Univ. Roma (1958) **17**, 423-438

Let us consider an n-dimensional random variable, represented on the Euclidean space S_n; the author studies the regression of an r-dimensional component variable (referred to a subspace S_r of S_n) on the complementary $(n-r)$-dimensional component variable (referred to the complementary space S_{n-r}), and particularly the *mean-square regression linear space* ("spazio lineare dei minimi quadrati") S^*_{n-r}, defined as the best linear estimate, according to the principles of least squares, of this regression.

The author calculates, by means of concise matrix algebra, the coefficients of the equations of S^*_{n-r} and then proves the following propositions:

1. the space S^*_{n-r} depends upon the space at infinity S^∞_{r-1} of S_r only; i.e. is invariant for the group of affine transformations under which S^∞_{r-1} does not vary;
2. confining ourselves to spaces through the mean point or referred to mean points as origin, the relation between S_r and S^*_{n-r} is reciprocal (i.e. the mean square regression space relative to S^*_{n-r} is S_r) and precisely is connected with the polar system of an ellipsoid of S_n.

(G. Dall'Aglio)

MOUSTAFA, M. D. (Institute of Statistics, Cairo University) **6.2 (6.1)**
A note on Partial Correlation—*In English*
Egyptian Statist. J. (1957) **1**, 1-5 (3 refs.)

The author considers the case of three correlated variables and studies the partial correlation coefficient between two of them.

He defines the partial correlation coefficient as the ordinary coefficient of correlation between the residuals obtained by subtracting from the two variables their true regression function on the third variable.

The form of the regression function is generally unknown and has to be assumed. The author proves that: If at least one of the regression functions is rightly assumed except for a constant, then the coefficient of correlation obtained by using the assumed regression function has the same sign and is in modulus less or equal to the true partial correlation coefficient.

In particular, if at least one of the regression functions is linear, he proves that the well known formula for the coefficient of partial correlation gives a value which has the same sign and is in absolute value less or equal to the true coefficient of partial correlation.

(J. B. Simaika)

RIZZI, A. (University of Roma) **6.0 (6.4)**

Cograduation tables in the three-dimensional case—*In Italian Statistica* (1959) **19**, 72-78 (4 refs., 8 tables, 1 fig.)

The purpose of this paper is to extend the cograduation tables, constructed by Salvemini, to the case of triple variates. Tables of this type occur in the calculus of the indexes of homophily.

In Italian usage, these indexes are special measures of concordance between statistical phenomena.

To overcome difficulties in regard to multiple variates, the author makes use of the technique used by Salvemini in connection with cograduation tables. In addition, he shows a logical scheme relative to a particular cograduation table. The plan can be adapted to automatic computation on electronic machines.

(V. Amato)

ROY, S. N. & GNANADESIKAN, R. (Institute of Statistics, North Carolina) **6.1 (4.8)**

A note on " Further contributions to multivariate confidence bounds "—*In English Biometrika* (1958) **45**, 581

The authors refer to a previous paper of theirs in *Biometrika* (1957) **44**, 399-410. In this previous paper dealing with the general linear hypothesis, confidence statements were derived and similar statements were obtained involving the truncation of columns of a given matrix.

The authors point out in this present note that there is another type of truncation which arises in discussing the uninvariate general linear hypothesis. They illustrate this by considering the problem of testing the equality of several treatment effects in ordinary analysis of variance.

(F. N. David)

SANKARAN, M. (Presidency College, Madras)
On Nair's transformation of the correlation coefficient—*In English Biometrika* (1958) **45**, 567-571 (5 refs., 2 tables) **6.2** (2.2)

Nair suggested a statistic $x = (r-\rho)/(1-\rho r)$ where r is the sample, and ρ the population correlation coefficients. It is shown that x and the inverse sine of x are both asymptotically normally distributed for n large and that the inverse sine is more quickly normal than is x. It is also shown that the sine of x can be used to evaluate the probability integral of r. A confidence interval for ρ is also obtained. The author remarks (p. 568) "... it is not clear that the inverse sine transformation of x has any practical advantage over the z-transformation, indeed it is not appropriate in some situations where z is useful."

(F. N. David)

SARGAN, J. D. (Leeds University)
The estimation of economic relationships using instrumental variables—*In English Econometrica* (1958) **26**, 393-415 (16 refs.) **6.6** (4.8)

This paper applies the use of instrumental variables, first suggested by Reiersøl in 1941 [*Econometrica* **9**, 1-23], to the more general case in which relationships are not exact: the ideal economic variables assumed generated by a set of dynamic stochastic relationships and the actual economic time series subject to random disturbances or measurement errors. The author considers the problem of determining the coefficients for a single relationship where the residuals are not autocorrelated. The main results are concerned with asymptotic properties of the estimates and to this end the author first considers the asymptotic covariance matrix. With this he proceeds to discuss the problem of choice between the possible instrumental variables in order to obtain the reduced set necessary to enable the method to be used. At the end of this section is a short discussion of the advantage of adding one new instrumental variable—stating that eliminating one from the reduced set cannot improve the asymptotic variance matrix.

In the fifth and sixth sections of his paper, the author draws a parallel between his analysis and the canonical correlation approach of Hotelling [*Biometrika* (1936) **28**, 321-335] and of an alternative derivation in the form of a minimax approach similar to that suggested by Durbin [*Rev. Int. Statist. Inst.* (1954) **22**, 23-54]. The paper then goes on to consider the asymptotic distribution of the smallest characteristic root and other matters prior to deriving two confidence regions for the unknown coefficients and a brief report of some lengthy (unpublished) calculations on the bias of the asymptotic approximation: the biases in the estimates tend to be large when the relationship is almost unidentified. The last two sections of the paper give a comparison with alternative methods of estimation: (i) Least Squares, and (ii) Limited-Information maximum-likelihood methods.

The author hopes subsequently to publish some of the large amount of work that has been carried out to test in practice the method outlined in this paper. From previous results and work some tentative conclusions are offered:

(i) The use of the instrumental variable method will produce convenient estimates of coefficients even when large measurement errors are apparent.

(ii) Better results may appear to be obtained if an attempt is made to reduce measurement errors—without introducing autocorrelation in the residuals.

(iii) Large measurement errors in predetermined variables are likely to cause biased estimates if the Limited-Information maximum-likelihood method is used. The Least Squares method also produces strongly biased estimates when there are large measurement errors, or when some of the independent variables are not predetermined and the relationship is confluent.

(iv) Large numbers of instrumental variables may not improve the accuracy of the estimates.

(Wm. R. Buckland)

SHEPS, M. C. (Harvard University Medical School) 6.9 (4.3)

An examination of some methods of comparing several rates or proportions—*In English* *Biometrics* (1959) **15**, 87-97 (10 refs., 4 tables)

The author considers various statistics used for quantitative comparisons of relative frequencies in the fields of public health and medicine. Studies like those of the association between smoking and lung cancer, between congenital malformations and exposure to radiation, or between coronary disease and obesity require some kind of quantitative measure of the " change in risk ". Rates, such as mortality rates, have been compared by calculating the ratio of two mortality rates, the ratio of two survival rates, and the difference between the two mortality or survival rates.

The author analyses these statistics in terms of several models. The model defined as Model 1 applies to the study of different populations with mutually independent risks. He maintains that for this model the difference in rates has an advantage over a ratio because the numerical value of the ratio depends on whether mortality rate or survival rate (one minus the mortality rate) is selected for comparison, while the difference is unaffected by this selection.

An alternative model (Model 2) is suggested for the common situation where the factor being investigated is assumed to modify the reaction of the population in one direction only. Under this model the expected rates in the several samples are not independent. As a result, the difference is a product of two probabilities, the ratio between one set of rates is compounded of several probabilities, while the ratio between the complementary set of rates is the estimate of a single probability. When a deleterious effect is expected, the ratio of success rates is an estimate of the diminished chance of " success " or survival ; and when a beneficial effect is expected, the diminished risk of failure is estimated by the ratio of " failure " or mortality rates.

The author derives the maximum likelihood estimates of the parameters, and their variances, for Model 2. The aforementioned statistics are then expressed in terms of these estimates. He proposes the use of a new statistic, an estimate of one of the parameters in Model 2. In terms of the smoking and lung cancer problem this parameter is the following probability of " risk " : Given that a smoker belongs to a category of persons who would otherwise escape lung cancer, he now suffers the risk that he will fail to do so.

This method is also applied to the poliomyelitis vaccination problem where the Model 2 parameter is the increase in protection given by vaccination rather than an increase in risk. The model is incapable of handling a situation that could occur in such studies, namely, of a vaccine actually producing the disease in some percentage of the cases.

Several different modifications of the model are indicated for the case with more than two samples.

(R. J. Taylor)

SIDDIQUI, M. M. (Univ. of N. Carolina ; now at Boulder Laboratories, Nat. Bur. Stand. U.S.A.) 6.1 (10.7)

Covariances of least-squares estimates when residuals are correlated—*In English* *Ann. Math. Statist.* (1958) **29**, 1251-1256 (4 refs.)

The author is concerned here with certain problems which arise in linear regressions when the residuals are not independent. Earlier work has given least squares estimates of the regression parameters and the residual variance, together with some asymptotic properties of these estimates ; references to this work are given. The contribution of this paper is to obtain explicit formulae for the elements of the covariance matrix and the expected value of the residual variance when a specific variance matrix is assumed for the residuals ; analytically simple bounds for the elements of the covariance matrix are also obtained. The residual variance matrix which has been assumed is one in which all elements an equal distance off the main diagonal are equal.

Two special cases have been worked out : (1) The true response is a linear combination of two non-stochastic variables, and an odd number of observations are made. (2) The true response is a linear combination of trigonometric functions which are themselves known functions of non-stochastic variables. The general formulae obtained earlier for the covariance matrix and expected value of the residual variance are specialised for each of these two cases.

This paper deals only with the situation where the residual variance matrix is known. It is pointed out that, while this is frequently not the case, the least squares estimates of the regression parameters and residual variance may still be obtained. Since this desirable property is not enjoyed by maximum-likelihood estimates, the author suggests the inadvisability of research in such a direction.

(D. H. Shaffer)

TINTNER, G. (Iowa State College) **6.7** (0.2)

A new method for estimating the logistic function—*In German* *Metrika* (1958) **1**, 154-157 (9 refs.)

Time series concerning developments in the field of economics or population, often show a trend which can be described in a satisfactory manner only by an exponential-function of the time t rather than by a polynomial in t. The fact that the polynomials become infinite for $t \to \infty$ makes them inappropriate especially for the purpose of extrapolation. For many purposes, the logistic function is well suited for describing the trend: $y_t = k/1+be^{-at}$. The methods commonly used for curve-fitting (method of least squares or maximum likelihood) lead to difficulties in this case, as the resulting equations for the unknown parameters are not linear.

A way out consists in considering the differential equation of the logistic function or the difference equation of its reciprocal $z_t = 1/y_t$. These equations are linear relationships between magnitudes which can be calculated, in principle, from the data originally given. The differential equation of the logistic function uses the rate of growth, $\frac{1}{y_t}\frac{dy_t}{dt}$, which, however, can only approximately be determined from the data given. The difference equation of $z_t = 1/y_t$, on the other side, is linear in z_{t+1} and z_t. Therefore, the common methods for estimating linear functions can be applied under the condition that the values of y_t are given for equidistant times. With this method the development of the population in Sweden is treated as an example. In treating the differential or difference equation instead of the original function, one parameter—the constant of integration—is lost. In the case of the logistic function, this parameter can be estimated by a formula of Rhodes.

(F. Ferschl)

WHITE, R. F. & GRACA, J. G. (Iowa State College, Ames) **6.8** (6.7)

Multinomially grouped response times for the quantal response bioassay—*In English* *Biometrics* (1958) **14**, 462-488 (15 refs., 3 tables)

In work with quantal response data experimental approaches may be classified according to the method of observation employed. Suppose doses $\xi_1, \xi_2, \ldots, \xi_m$ are administered to $n_1, n_2, \ldots, n_m$ individuals respectively. Let the number of individuals responding to dose ξ_i be observed at times $\tau_{i1} < \tau_{i2} < \ldots < \tau_{ik_i}$, which are pre-specified. The data are then said to be multinomially grouped according to response times.

The aim of the authors' analysis is to estimate a relation between p, ξ, and τ, where p is the proportion of the population that responds to an amount of stimulus ξ in a period of time τ. Particular attention is given to relations for which a transform y of p is a linear function (additive model) of transforms x and t of ξ and τ, or a linear function plus an interaction term (non-additive model). A detailed discussion is given for these two models when y is the logit transform of p. Considerable attention is given to the problem of choosing metameters. Two procedures for choosing these are presented for the above particular models, and an example is given in which the dose metameter is log dose and the time metameter is the reciprocal of the square root of time. The troublesome case when no response is obtained in some time intervals is discussed and an adjustment is recommended.

Since the response numbers are accumulated for the successive times within a dose, the observed proportions are not independent within doses. This introduces some complexity into the analysis which is at least partly resolved by use of the minimum modified chi-square method of estimation of the coefficients of the relationship. Computational details are given for this method of estimation. An appendix contains the derivation of the estimators of constants in the functional form by both the method of maximum likelihood and the method of minimum modified chi-square. A derivation of the variances of the estimates obtained from the estimated relation is given in the appendix for the additive and non-additive models.

(C. P. Quesenberry)

COX, C. P. (National Institute for Research in Dairying, Univ. of Reading) **7.6 (9.7)**

Experiments with two treatments per experimental unit in the presence of an individual covariate—*In English*

Biometrics (1958) **14**, 499-512 (13 refs., 9 tables)

The methods available for eliminating temporal or spatial trends in experimental material from treatment comparisons include stratification in the design and covariance on the trend variable. They eliminate the mean trend over a set of experimental units, and are likely to be inadequate in some situations, particularly when within-block treatment comparisons are desired. The author has previously given a method for removing individual curvatures and isolating treatment comparisons in Latin square designs. In the present work this design restriction is removed, and a general solution obtained, giving within-individual treatment comparisons free from general curvature effects (assumed independent of treatment effects). The case considered is that in which there are two treatments per individual or block. Time trends are removed by the use of equally spaced observations. Upon making necessary changes, one may render the solution applicable to the more general problem of unequally spaced observations.

For each individual the time is divided into three periods consisting respectively of t_1, s, and t_2 equal and successive intervals. The two treatments, of which the main effects are to be compared, are administered, one during each of the first and last periods. One independent observation is obtained in each interval, yielding t_1 observations in the first period and t_2 in the third. No observations are required in the intervening period of s intervals. In biological work this may serve as a conditioning period; in some cases it may be possible to take $s = 0$.

For the complete experiment, a mixed model is used with fixed treatment effects and random individual and interaction effects. In application to an individual the treatment and interaction effects are confounded. The least squares technique is used to isolate a comparison involving the difference between the two treatment effects and the difference of corresponding interaction effects. The sum of squares for treatments and interactions, as well as those for regressions up to the desired order are obtained. The establishment of orthogonality properties is indicated and the analysis of variance set up.

Tests of significance in the complete experiment, are discussed and the precision of the method is examined. Reference is made to the question of allocating values to t_1 and t_2 so as to minimise the contribution of the unexplained variation to the variance affecting the treatment comparisons. A solution is given for the case of linear gradients, explicit solutions for other cases not being as yet available. Consideration is given to the testing of the adequacy of the mathematical model and the examination of the validity of the associated assumptions. The procedure is illustrated in a numerical example.

(W. A. Glenn)

HALL, W. J. (University of North Carolina, Chapel Hill) **7.7 (1.8)**

Most economical decision rules—*In English*

Ann. Math. Statist. (1958) **29**, 1079-1094 (15 refs., 1 fig.)

The problem considered is to determine optimum decision rules for what is often referred to as the discrimination problem, i.e. the problem of choosing one of m mutually exclusive alternative hypotheses A_i. The decision procedure then is to subdivide the sample space into m mutually exclusive regions R_i and if the sample point falls in R_i, the decision A_i is made (or hypothesis A_i accepted). The author denotes by $p_{ij}(D)$ the probability that the decision procedure leads to decision A_i when in fact A_j is true. The matrix $\| p_{ij}(D) \|$ characterises a decision rule or discriminant function.

The author then gives two definitions of a most economical decision rule. A decision rule D_N is most economical with respect to the vector $\alpha = (\alpha_1, \alpha_2, \ldots \alpha_n)$ if

$$p_{ii}(D_N) \geqq \alpha_i \; (i = 1, 2, \ldots n)$$

and N is the least integer n for which this inequality is achieved by any decision rules based on samples of size n. In other words such a decision ensures that the probability of rejecting A_i when A_i is true has prescribed bounds and it does so for the smallest possible sample size.

A decision rule D_N is most economical with respect to the matrix $\| \beta \|$ if $p_{ij}(D) \leqq \beta_{ij}$ $(i, j = 1 \ldots n)$ and N is the least integer n for which this inequality is achieved by any decision rule. These two definitions are extensions of the Neyman-Pearson Type 1 and Type 2 Error concepts respectively.

However, the author then sets up special types of loss functions, which lead to a solution of the problem of optimum discriminant functions from results of Wald's minimax theory. Conditions are given that a most economical decision rule exists. These are given both in fairly general terms and for families of distributions of exponential type.

Some examples are given where the random variables are normal and the hypothesis A_i specifies that their mean lies in some interval (θ_{i-1}, θ_i). As is intuitively to be expected, the decision regions are of the form accept A_i, if $c_{i-1} \leqq \bar{x}_i < c_i$ (σ known) or $c'_{i-1} \leqq t < c'$ (σ unknown) where t is the usual Student ratio. The c's are to be found by solving iteratively a series of equations in $n+1$ unknowns involving the normal distribution function. The equations for the c''s are similar except that the normal distribution is replaced by the non-central t-distribution. Since a different set of these equations must be solved for each α (or $\| \beta \|$) and set of θ's, the exact optimum solution is not too well suited as yet for practical use.

(D. G. Chapman)

MATHER, K. & JONES, R. MORLEY (University of Birmingham) **7.3** (–.–)
Interaction of genotype and environment in continuous variation : II. Analysis—*In English Biometrics* (1958) **14**, 489-498 (4 refs., 2 tables)

This paper examines the requirements for successful separation of the effects of genotype-environment interaction from other components of variation, and for the estimation of these components by comparisons and analyses of the variances and covariances which are obtainable in experiments.

When polygenically determined continuous variation is involved, important comparisons can be achieved by raising true-breeding lines and their F_1 in the same range of environment where these environments are so distinguished from one another that differences between them are large in their effects when compared with variation within them. Then genetic parameters can be estimated as well as the environment and interaction parameters.

When statements of genetical constitution must be based on knowledge of ancestry, on observations or on assumptions of linkage and selection, as with continuous variation, then families must be grown over a comparable range of environment. If the range of environments is large enough so that the sampling variation is negligible in respect of genotype and of environments then the differences between their mean phenotypes will reflect only their differences in average genetical constitution. Variation between individuals will reflect the differences amongst the environments of individuals and also interactions. If the families are distributed at random over the environment, this paper furnishes the variances of the segregating generations.

If the individuals of all families are not distributed at random over a common range of environments, but if each family is raised in one or more compact groups, which are themselves distributed over a wider range of environments, the environment of differences affecting the variation of individuals within the plot are then different from those affecting the variation in plot and family means. In this case a method is given to separate the effects of environmental differences into those within plots and those between plot means. The interaction components of variation between plots may be isolated, but analysis of variation within plots is again possible only in terms of transforms.

A previous paper, Part I—Description, was published by the authors [*Biometrics*, **14**, 343-359].

(C. Y. Kramer)

DARROCH, J. N. (Department of Mathematics, University of Cape Town) **8.3** (4.3)
The multiple recapture census. I. estimation of a closed population—*In English Biometrika* (1958) **45**, 343-359 (16 refs.)

In this paper the author considered multiple-recapture census for which the number of samples s is usually considered fixed. In a few cases s is allowed to be variable. The sample sizes are regarded both as constants, and as binomial variables. The investigation is restricted to a closed population in which there is neither departure due to death or emigration, nor augmentation due to birth or immigration. The author will take account of both departure and augmentation in a second paper.

Two alternative models are constructed: Model (A), a multinomial density and claimed to be new, is applicable when the effort put into the catching of every sample is fixed before the experimenter starts; Model (B), a generalised hypergeometric density which forms the basis of most of the previous investigations, is used when the experimenter is determined to catch a specified number of individuals at each sample.

The two models are shown to lead to the same estimate of the population, n, and to the same asymptotic variance of this estimate. Model (A) is introduced as it is capable of generalisation which (B) is unable to accommodate. Confidence interval estimation under Model (B) and a special case of one individual per capture are investigated.

In the case where s is a random variable, a model (A_1) is constructed and finally the author proves that for a certain class of estimates of n, those derived form (A) and (A_1) by the method of maximum likelihood are asymptotically the best.

(C. Crouse)

DE LUCIA, L. (University of Rome) **8.1** (4.4)
Problems of sampling techniques in market research—estimation of proportions—*In Italian Quad. Ass. Ital. Studi Mercato*, no. 3, Giuffré, Milano (1958) (19 refs., 12 tables)

In this report of 128 pages are collected the lessons held, during a course in Market Research, on the estimation of proportions in random sampling of attributes, which, as the author puts in evidence, is one of the most frequent cases in Marketing problems. The author exposes in an elementary form the basic notions of this field in the case of random sampling without replacement, emphasising the role of the problem of repeated trials and of the connected random variables.

The first part is devoted to simple sampling. The author exposes first the classical methods and formulas for estimating the proportion and the difference between two proportions, and for evaluating their standard errors. Then, reporting some results given in a previous paper [*Riv. Ital. di Econ., Demog. e Stat.* (1957) **11**, 123-142], he deals with the estimate of proportions with regard to a part of the universe. In this study, as he shows, it is possible to include the case of a sample in which are present sampling units that cannot be classified. In the last section the methods of estimating the sampling variance by means of the sampling data are given.

In the second part, devoted to stratified samples, the author repeats the topics of the first one. There is added the treatment of two most usual methods of dividing the total sample size among the strata: the partition being proportional to strata sizes, together with optimum allocation. Some fictitious examples, extensively worked out, accompany the development of the theory.

(G. Dall'Aglio)

ÉLTETŐ, Ö. & SARKADI, K. (Math. Inst., Hungarian Academy of Sciences, Budapest) **8.8** (1.8)
On the statistical problems arising in microbiological examination of tomato purées—*In Hungarian*
Mag. Tud. Akad. Kut. Mat. Intézet. Közleményei (1958) 219-226 (2 refs., 2 tables)

The authors deal with questions connected with the following problem :

A lot of tomato purée will be accepted by the consumer if the Howard-count found in the sample taken out from the lot in question does not exceed a fixed limit. What limit is to be used by the producer on the occasion of a previous quality control for the purpose of retaining the bad lots ? It is assumed to be legitimate to regard the distribution of the Howard-count as a Normal one. The variance is assumed to be known from empirical data. The error arising from the heteroscedasticity can be eliminated by a homoscedastic transformation.

The problem is discussed according to the following three principles :

1. It is prescribed that the joint probability of accepting in the first and rejecting in the second examination should not be greater than $1-\beta$. The solution is given by formula :

$$x = t+2\sigma\Phi^{-1}(\sqrt{1-\beta}).$$

σ^2 denotes the variance of the Howard-count, t and x the critical value in the final and in the preliminary sampling, respectively, and $\Phi^{-1}(u)$ denotes the inverse of the standardised normal distribution function

$$\Phi(u) = (2\pi)^{-\frac{1}{2}}\int_{-\infty}^{u} e^{-z^2/2}\,dz.$$

2. Let w_1 be the loss caused by the wrong decision in the first examination if we retain a lot which would be accepted, and w_2 the loss if we dispatch a lot which will be rejected. In this case the minimax solution is given by formula

$$x = t+2\sigma\Phi^{-1}\left(\frac{w_1}{w_1+w_2}\right).$$

3. It is assumed that the expected value of the Howard-count of a lot is itself a random variable varying in the lots transported from the same producer according to the normal probability law. The problem can be solved numerically by the aid of the table of the two-dimensional normal distribution.

(I. Vincze)

GARNER, N. R. (Aerojet-General Corporation, Azura, California) **8.8** (5.7)
Curtailed sampling for variables—*In English*
J. Amer. Statist. Ass. (1958) **53**, 862-867 (6 refs., 1 table)

The usefulness of curtailed sampling is evident when testing is costly, when each test takes a long time, or when failure might constitute a hazard. Thus, if under a single sampling plan specifying sample size n the eventual rejection decision can be foretold with certainty after only $r<n$ observations, it will be advantageous not to sample the remaining $(n-r)$ items. Curtailed sampling has long been prescribed for attribute single-sampling ; for variables single-sampling, curtailed procedures have been suggested in the past but no optimum condition for this kind of sampling procedure has as yet been suggested.

This paper supplies such an optimum condition for curtailed sampling of variables in the one-sided case. The condition is optimum in the sense that, if sampling is curtailed and the " reject " decision made after only $r<n$ observations, then there exist $(n-r)$ possible remaining observations such that rejection would " just barely " be indicated on the basis of the full sample of n.

Suppose the lot is to be rejected if $\bar{X}+ks>L_u$, where $\bar{X}$ and s are the mean and standard deviation for the entire sample of n observations and k and L_u are specified by the sampling plan. Then curtail the sampling and reject if $\bar{X}_r+k_rs_r>L_u$, where $\bar{X}_r$ and s_r are the mean and standard deviation for the curtailed sample of the first r observations and $k_r = k\sqrt{\frac{r-1}{n-1}-\frac{(n-r)(r-1)}{nrk^2}}$.

If the lot is to be rejected when $\bar{X}-ks<L_l$, curtail and reject if $\bar{X}_r-k_rs_r<L_l$.

(H. T. David)

GHOSH, S. P. (Calcutta University) **8.1** (8.7)

A note on stratified random sampling with multiple characters—*In English*
Bull. Calcutta Statist. Ass. (1958) **8**, 81-90 (3 refs., 5 tables)

The optimum allocation in a case of stratified random sampling for multiple characters is defined in this note as that allocation which for a fixed total sample size minimises the generalised variance, namely, the determinant of the variance covariance matrix (σ_{ij}) of the estimates of the population values. The equations providing the optimum values of strata allocations as functions of all the variances and covariances are complicated. However, for obtaining the solutions numerically, an iterative procedure is suggested. As a first approximation, the individual minimum values of each σ_{ij} is obtained and substituted in a set of equations to give the values of n_i for the different strata. The values of n_i, so obtained, were substituted in another set of equations to obtain the second set of approximations of σ_{ij}. This process is repeated till stable values of n_i are obtained. The condition under which the proposed optimum allocation will be identical with the average of Neyman's allocations for the different characters is also considered.

A numerical example with six strata and three characters is used to illustrate the proposed allocation with a view to demonstrate empirically the rapidity of the convergence of the iterative process. Optimum strata allocations have been obtained

(i) assuming the correlation between two characters to be the same for all strata and
(ii) assuming that there is no correlation between the characters.

Using this example, the variances of the estimates of the population means under the optimum allocation have been compared with those obtained by

(i) proportional allocation,
(ii) average of Neyman's allocation, and
(iii) allocation arrived at by minimising the sum of the ratios of the variances for the characters to their respective optimum variances.

(M. N. Murthy)

HAMMOND, L. C., PRITCHETT, W. L. (Univ. of Florida, Gainesville) **& CHEW, V.** (North Carolina State College, Raleigh) **8.1** (8.7)

Soil Sampling in Relation to Soil Heterogeneity—*In English*
Proc. Soil Sci. Soc. Amer. (1958) **22**, 548-552 (12 refs., 10 tables)

After reviewing recent statistical studies that have been made on the problems of soil sampling, the paper presents analyses of soil samples taken from cultivated fields (tobacco and corn), pastures and orchards (oranges). The seven locations ranged in size from 2 to 35 acres and included 6 representative soil types in Florida. The samples were analysed for moisture, *p*H, and extractable potassium, calcium, magnesium and phosphorus.

A three-stage sampling scheme was adopted. The number of first-stage units in the field and in the sample varied with the size of the field, but the number of second-stage units was 9 in all cases, of which 2 were randomly selected. Two one-inch diameter cores were taken from each second-stage unit. There was no compositing of cores.

Applying the usual analysis of variance techniques, but taking the finiteness of the populations of sampling units into consideration, variance components were calculated for each of the fields and for each of the variables measured. The efficiency of the multi-stage sampling scheme relative to simple random sampling was found to be high (generally over 70 per cent.) for "macrouniform" fields, i.e. for fields where the variation among the larger units is small compared to variation among the subunits. For such fields, the administratively more convenient multi-stage sampling is recommended over simple random sampling.

The paper concludes with some cost consideration and optimum allocation of the sampling units. For a simple assumed cost function, the larger number of samples required for multi-stage sampling scheme to achieve a given precision actually cost only about one-third that for simple random sampling.

(V. Chew)

MOSHMAN, J. (Corp. for Econ. and Indust. Res., Arlington, U.S.A.) **8.3** (–.–)

A method for selecting the size of the initial sample in Stein's two-sample procedure—*In English*

Ann. Math. Statist. (1958) **29**, 1271-1275 (5 refs., 1 table)

In 1945, Stein published a paper [*Ann. Math. Statist.*, **16**, 243] giving a two-sample procedure whose power was independent of the variance. No specification was given for the choice of the size of the initial sample but in 1953 a paper by Seelbinder [*Ann. Math. Statist.*, **24**, 640] suggested that it should be selected to minimise the expectation of the total sample. An efficient choice will hold the expectation of the sample small and will further reduce the probability of an extremely large total sample. This note by the present author explores the problem and shows that an upper percentage point of the distribution of the total sample size, when used in conjunction with the expectation, is a guide to the efficient choice of size for the initial sample. The procedure outlined is a rapidly calculable function based upon ordinary χ^2 (distribution) tables and is a function of the population variance.

If an initial sample (N_0) is taken and an estimate s^2 (based upon N_0-1 degrees of freedom) of the population variance is computed, then the total sample size, N, is: max $[(s^2/z)+1, N_0]$ where $z>0$. It follows that $P(N=N_0) = P(ns^2/\sigma^2) = P[\chi^2/n \leqslant nN_0z/\sigma^2]$ where $n=N_0-1$. Similarly $P(N=m), = P[n(m-1)z/\sigma^2 < \chi^2/n \leqslant nmz/\sigma^2]$. These two expressions facilitate the calculation of a given percentage point, N_p, of the distribution of N: in fact,
$N_p = \max\{N_0, \lambda^{-1}[100\ p\text{th percentage point of } \chi^2(n)/n]+1\}$
where $\lambda = z/\sigma^2$.

An objective rule is given for the choice of N_0 and it is suggested that p would be chosen independently from non-statistical grounds. An example is given ($\lambda = 0{\cdot}1$) together with a brief discussion of some subsidiary questions.

(Wm. R. Buckland)

STANGE, K. (Technische Universität Berlin) **8.8** (11.2)

The graphic treatment of plans for quantitative inspection—*In German*

Metrika (1958) **1**, 111-129 (3 refs., 10 figs.)

The paper deals with sampling procedures for variables. The use of variables inspection (measurement) instead of attribute inspection (classification by gauges) is of advantage, if the cost of measurement is nearly equal to the cost of classification, if the cost of inspection is very high anyway (for example due to the destruction of the inspected items) or if the fraction defective is very small.

The quality characteristic x is assumed to be normally distributed with parameters μ, σ. An item is regarded to be defective if x exceeds some given value T. The quality of a lot is judged by calculating the test statistic $z = \bar{x}+ks$ and accepting the lot if $z \leqslant T$, rejecting if $z>T$. z has a mean of $\xi = \mu+k\sigma$ and variance

$$\sigma_2 = \frac{\sigma^2}{n}\left(1+\frac{k^2}{a}\right)$$

(n = sample size). It is approximately normally distributed for $n \geqslant 10$. In order to obtain the acceptance probability, one has to consider $u_W = T-\xi/\sigma^2$, which is a linear function of the magnitude $u = T-\mu/\sigma$ which is relevant for the fraction of "defective" items (items with $x>T$). The coefficients of this linear function depend only on n and k. Therefore, the operating characteristic (*OC*) curve, giving the acceptance probability as a function of the fraction defective, is a straight line if plotted in double-probability paper (i.e., a probability paper, which has probability scales on both axes). Therefore, if two points of the *OC* curve are given (for example producer's safe point, producer's risk and consumer's safe point and consumer's risk), it can be plotted easily on this double-probability paper. Starting from the *OC* curve on the double-probability paper, the sample size n and the constant k can easily be determined by graphical methods. The use of this graphical method makes the use of any table superfluous.

In order to condense the information, given by a sequence of samples (for instance in order to decide whether the inspection level should be reduced or tightened), usually the statistic

$$t' = \frac{T-\bar{x}'}{s'}\left(\bar{x}' = \frac{1}{m}\sum_{i=1}^{n}\bar{x}_i,\ s'^2 = \frac{n-1}{mn-1}\sum_{i=1}^{m}s_i^2\right)$$

is used. It is proposed to use instead of t' the statistic

$$\bar{t} = \frac{1}{m}\sum_{i=1}^{m}t_i \quad \left(t_i = \frac{T-\bar{x}_i}{s_i}\right).$$

Both, t' and $\bar{t}$, are asymptotically normally distributed with the same variance and therefore equally well suited for this purpose. $\bar{t}$ can, however, be easily determined by graphical methods, as the magnitudes t_i are already given.

Similar results concerning attribute inspection will be given in a future paper. In this paper, also a simple equation will be given for the relation n_M/n_N which compares the sample sizes in the case of measurement (n_M) and classification (n_L).

(J. Pfanzagl)

STANGE, K. (Technische Universität, Berlin) 8.9 (10.7)

Belt Sampling—*In German*

Metrika (1958) **1**, 177-223 (5 refs., 20 figs.)

Many production processes, especially in the chemical industry, are of a continuous nature. For supply and delivery a belt system is used. If samples are taken from the belt, in many cases the results of subsequent samples are correlated. The correlation is a decreasing function of the distance between the samples. The shortest distance for which adjacent samples are non-correlated is called the region of correlation.

The accuracy of the sample mean $\bar{x}$ depends on the variance of the individual values, the sample size n, the way in which the sample is taken (random sample, stratified sample or systematic sample), the length of the belt L, the distance between two adjacent samples l and the region of correlation λ. If $l>\lambda$ the correlation is of minor importance and the variance calculated from the sample can be used for estimating the variance of $\bar{x}$. If $l<\lambda$ the correlation may be of great importance. A correlation integral is introduced, by means of which a critical distance of adjacent samples and a critical sample size n_* is defined.

If the quality of a given belt-section of length L is to be inspected by systematic sampling (i.e. sampling at equal distances), the variance decreases very rapidly (in some instances to the order of $1/n^2$) if the sample size is greater than n_*. The variance within the belt section can practically be eliminated by taking a sample of size $2n_*$.

If the purpose of the inspection is to judge the process itself, samples of size L are taken from different belt sections. In this case the variance converges for $n\to\infty$ to the value of the correlation integral. For the correlation functions usually observed this limiting value is reached for $n = 2n_*$. Therefore the accuracy cannot be increased by taking samples larger than $n = 2n_*$. Thus in both cases there is no essential gain in accuracy by taking samples of a size greater than $2n_*$.

These results also hold true if the subject under investigation are mean values for longer periods. If on the other hand the purpose is to regulate the process, values for specific time points are needed. In this case the accuracy can be increased by taking samples at shorter distances.

The results of this inquiry are of practical importance only if the correlation function is stable during time. Whether this is the case can not yet be decided and many empirical investigations will be necessary to clarify this question.

(A. Adam)

STEVENS, W. L. (University of São Paulo, Brazil) 8.1 (4.3)

Sampling Without Replacement with Probability Proportional to Size—*In English*

J. R. Statist. Soc. B (1958) **20**, 393-397 (1 ref.)

This paper puts forward an alternative to Yates & Grundy's [*J. R. Statist. Soc.* B., 1953] extension of Horvitz and Thompson's [*J. Amer. Statist. Assoc.*, 1952] result and it is not self-explanatory without reference to both of the earlier papers. A finite population of N units, each of which possesses a measurable characteristic y, is supposed. It is desired to estimate the mean value (μ) of this characteristic in the finite population using the y-values of a sample of N units. In drawing the sample the probability of selecting a particular unit is proportional to a second measure denoted by x which measures the size of the unit. The x-measures of each and every member of the finite population are supposed known.

Stevens considers the case where the elements of the population fall into groups having the same size, there being N_i elements of common size x_i in the ith group and $N_i \geqq n$ for all i. He proposes that n groups be selected with replacement, the probability of selection of the ith group being proportional to the total sizes, i.e., $N_i x_i$, in the group and that if the ith group is selected t_i times then t_i units are selected from that group without replacement. On this basis he obtains an unbiased estimator called $\bar{r}$, of the total $N\mu$ which is formally the same as that for the case of sampling with replacement. He also obtains the variance of this estimator and an unbiased estimator of this variance.

Stevens considers that Yates & Grundy's method is "hardly practicable" for $n>2$ but that his method can be easily applied for an n of (say) ten with, if necessary, certain modifications (e.g., pooling the size groups) which he discussed and claims will not materially affect his results.

If p_i is the probability of selecting any given element in the ith group and if r denotes the result of dividing the y-value of an element by its probability of selection (viz. p_i if its size-group is the ith) then $\bar{r}$ is the mean of the n sample values of r. Since the probability of selecting t elements in the ith group is the binomial probability of selecting t in a sample of n independent events with constant probability $N_i p_i$, and since the conditional probability of selecting a given element from amongst those in the ith group is t/N_i, then the overall probability of selecting that element is np_i. Hence, $E(n\bar{r}) = nN\mu$. The occurrence of $n\bar{r}$, called $\sigma^2(R)$, is obtained in a similar but more lengthy manner by summing the conditional probabilities over the multinomial probabilities of selecting given numbers of units from a specified set of groups, thus obtaining the probabilities of selecting a given pair of elements. The unbiased estimator of this variance is simply obtained by observing that the mean value of the sum of squares of the n r-values about $\bar{r}$ differs from $(n-1)\sigma^2(R)/n$ by a simple multiple of $\Sigma N_i\sigma_i^2$ (where σ_i^2 is the variance of the y-values within the ith group). An unbiased estimator of this is found by means of a weighted sum of the within-group variances of the r-values.

(D. E. Barton)

WEILER, H. (Comm. Sci. Ind. Res. Organisation, Prospect, N.S.Wales) 8.9 (2.7)

Mean charts of constant false alarm risk for non-normal distributions—*In English*

Aust. J. Appl. Sci. (1958) **9**, 326-331 (5 refs., 4 figs.)

In quality control methods, the mean μ of a population is usually controlled by entering the means $\bar{x}$ of random samples of fixed size n into a control chart with limits $\mu \pm B\sigma/\sqrt{n}$, where σ is the standard deviation of the population and B = 3 or 3·09. If the population is normal, the probability that $\bar{x}$ falls outside these limits is 0·001 when $B = 3{\cdot}09$, so that such an event mostly indicates a change in population mean. However, many populations occurring in practice are non-normal, so that *false alarms* ($\bar{x}$ outside the limits for unchanged μ) may occur more often than expected on the basis of normality.

In this paper, the author establishes a relation between n and B such that control charts will provide the same protection against false alarm when non-normal distributions are encountered. It is shown how the width of the control limits can be much reduced as the size n of the sample increases.

By using an expansion in terms of the standard normal distribution $\phi(x)$ and its derivatives for the probability p that $\bar{x} > \mu + \beta\sigma/\sqrt{n}$, an expression is obtained for the *false alarm risk* $p_1 = p_0 + \Lambda\phi''(B)/6\sqrt{n}$, an upper barrier to the values of p for a given set of populations whose skewness does not too often exceed Λ. Here p_0 is the false alarm risk for a standard normal population, and $\phi''(B)$ the second derivative of the standard normal distribution at $x = B$.

Curves of constant false alarm risk per sample are given, in which B is plotted against n for fixed values of $\Lambda = 0$, 0·5, 1, 1·5, 2. Further curves of constant false alarm risk $p_1 = 0{\cdot}001$, 0·002, 0·003, 0·004 per sample, with B plotted against n are provided for $\Lambda = 1$, 2.

Lastly, curves of constant false alarm risk per article tested are given, B being plotted against n for $\Lambda = 0$, 0·5, 1, 2. The average number of articles tested before raising an alarm is found to be about 4000, 1877, 1176, 690 respectively in each of these cases.

(J. Gani)

ZINGER, A. & ST. PIERRE, J. (University of Montreal, Canada) 8.7 (1.8)

On the choice of the best amongst three normal populations with known variances—*In English*

Biometrika (1958) **45**, 436-446 (9 refs., 4 tables, 2 figs.)

This is a problem in statistical decision functions. It is given that there are three normal populations with known variances. Decisions have to be taken in (i) the problem of selecting the population with the largest mean given a probability of making a wrong decision, and in (ii) the problem of determining the smallest sample sizes required to select the population with the largest mean—for given values of differences between population means and for given values of decision probabilities. The statistic which is used is the standardised difference between the two largest sample means. Tables whereby the significance of the statistic under (i) can be judged are not given but can be obtained from the author if desired. An alternative procedure based on the difference between the largest sample mean and the grand mean is compared with the procedure proposed under (i) and it is suggested that this alternative procedure is not so good.

Under (ii), the planning of an experiment, some assumption is necessary about differences between the population means. An abacus is given whereby, for fixed values of the first and second kind of error, the equations of condition for the sample sizes can be solved. This solution was proposed by Bechhofer: an optimum solution is achieved by repeated application of Bechhofer's method.

(F. N. David)

ATIQULLAH, M. (University of Dacca, Pakistan) 9.1 (-.-)

On configurations and non-isomorphism of some incomplete block designs—*In English*
Sankhyā (1958) **20**, 227-248 (20 refs.)

This paper discusses the existence and isomorphism of, and relations between, certain classes of Balanced Incomplete Block (BIB) and Partially Balanced Incomplete Block (PBIB) designs. Let W_n denote the class of balanced incomplete block designs with parameters $v = n(n-1)/2$, $b = n(n+1)/2$, $k = n-1$, $r = n+1$, $\lambda = 2$ and π_n the class of two associate partially balanced incomplete block designs with parameters $v = b = n(n-1)/2$, $r = k = n-1$, $\lambda_1 = 2$, $\lambda_2 = 1$, $n_1 = (n-2)(n-3)/2$, $n_2 = 2n-4$, $p^1_{11} = (n-4)(n-5)/2$. The following theorems are proved:

Theorem 1: A necessary and sufficient condition for W_n to exist for $n \geqq 4$ is that π_n exists.

Theorem 2: If $n \geqq 4$ and $n(n-3)/2$ odd, a necessary condition for π_n to exist is that $(n-1)$ is a perfect square.

The non-existence of designs W_n for $n = 6, 9, 13, 14$, and $n = 18$ follows from these two theorems.

The problem of obtaining all non-isomorphic solutions of ϕn for various values of n is then taken up. It is shown that for $n = 4$, just one and for $n = 5$, just five solutions exist: for $n = 8$ two new solutions are obtained.

It is further shown that a Triangular Singly Linked Block design Γ_n which is a two-associate PBIB design with parameters $U = n(n-1)/2$, $b = n$, $r = 2$, $k = (n-1)$, $\lambda_1 = 1$, $\lambda_2 = 0$, $n_1 = 2n-4$, $n_2 = (n-2)(n-3)/2$ and $p^1_{11} = n-2$ has a unique solution for any n.

Lastly, two methods of generating two-associate PBIB designs D_n with parameters $v = n(n+1)/2$, $b = n(n-1)/2$, $k = n+1$, $r = n-1$, $\lambda_1 = 2$, $\lambda_2 = 1$, $n_1 = (n-1)(n-2)/2$, $n_2 = 2n-2$, $p^1_{11} = (n-3)(n-4)/2$ from a symmetric BIB design with $r = n+1$ and $\lambda = 2$ are described and it is shown that the two methods may lead to non-isomorphic solutions for D_n.

(J. Roy)

FREEMAN, G. H. (East Malling Research Station, Maidstone) 9.1 (9.7)

Families of designs for two successive experiments—*In English*
Ann. Math. Statist. (1958) **29**, 1063-1078 (9 refs., 7 tables)

The author discusses all known families of $O : PP$ designs with two associate classes, where a design of type $O : PP$ is one in which the first set of treatments is arranged in randomised blocks (a type O, for orthogonal, design) and the second set of treatments is arranged as a partially balanced incomplete block design with two associate classes (a type P design) with respect to both the blocks and the first set of treatments (referred to as rows and columns, respectively).

The $O : PP$ designs are classified according to the association scheme of the P-type design since, for practical purposes, the $O : PP$ designs have been constrained to have the same associate classes with respect to rows and columns. The majority of the $O : PP$ designs are group divisible with the remainder having triangular and Latin square association schemes. The author states that it appears unlikely that any other association schemes will provide $O : PP$ designs.

The author shows that not every pair of group divisible P-type designs can give rise to an $O : PP$ design. However, there are 56 pairs that do and the parameters for the resulting 56 families of group divisible $O : PP$ designs are listed. Also listed are the parameters for one family of triangular designs and one family of Latin square designs. The numerical values of the parameters are tabulated for all the "useful" designs (designs having more than 2 replicates or treatments, not more than 30 replicates, treatments, rows or columns and not more than 150 plots in all). Most of the useful group divisible designs occur in 3 of the 56 families, namely, certain subsets of the designs which are either singular with respect to both rows and columns, or singular with respect to one and either regular or semi-regular with respect to the other. The remainder of the useful group divisible designs are confined to 10 of the remaining 53 families. For each of the 43 families of group divisible designs that have no useful designs the author has tabulated the smallest designs that are theoretically possible. In all, the author has listed the parameters of 127 useful group divisible designs, eight triangular designs and two Latin square designs.

A catalogue of the useful group divisible and triangular designs has been prepared and is available at East Malling Research Station. General methods of analysis of the $O : PP$ designs have been given by the same author in one of the references.

(John W. Wilkinson)

GOMES, F. P. & GUIMARÃES, R. F. (São Paulo Univ. & Comp. Paulista de Estradas de Ferro, Brazil) **9.7** (9.1)
Joint analysis of experiments in complete randomised blocks with some common treatments—*In English*
Biometrics (1958) **14**, 521-526 (6 refs., 4 tables)

The paper deals with the intra-block analysis of a group of experiments in complete randomised blocks, where some treatments, but not all, are common to the whole group. The common treatments are called by that name and the others are called " regular treatments ". The " regular treatments " belonging to one experiment are not common to any other experiment.

The design studied for the analysis is a special case of intra- and inter-group balanced incomplete blocks, proposed earlier by Rao in a paper listed in the references. The whole set of experiments is regarded as one trial with incomplete blocks, and under the assumption of equal variances a joint analysis is carried out.

Least squares estimates of the treatment means on Model I are obtained. For the common treatments these are simply the arithmetic means ; however, the estimate of the mean of a regular treatment in any experiment has to be adjusted by subtracting from the arithmetic mean the difference between the mean of the common treatments in that experiment and the overall mean of the common treatments.

The estimates of the variances of various possible contrasts among the estimated treatment means are also given. These variance estimates are

(1) $\hat{V}(\hat{Y}_1)$ of the estimated contrast $\hat{Y}_1$ between the means of two regular treatments in the same experiment.

(2) $\hat{V}(\hat{Y}_2)$ of the estimated contrast $\hat{Y}_2$ between the means of two regular treatments in different experiments.

(3) $\hat{V}(\hat{Y}_3)$ of the estimated contrast $\hat{Y}_3$ between the means of two common treatments based on all observations.

(4) $\hat{V}(\hat{Y}_4)$ of the estimated contrast $\hat{Y}_4$ between the means of a common and a regular treatment.

In the case whan only one treatment is common $\hat{V}(\hat{Y}_1) = \hat{V}(\hat{Y}_4)$ and the contrasts $\hat{Y}_1$ and $\hat{Y}_4$ are similar. As a test of these contrasts Tukey's studentised range test is proposed with k = no. of regular treatments + 1. The other type of contrast which may be considered is that of (2). Here again Tukey's test may be used with k = number of regular treatments. When only contrasts between the common treatment and a regular one are to be considered, Dunnett's procedure with least significant difference = $t\sqrt{\hat{V}(\hat{Y}_4)}$, where t refers to the multivariate Student t-distribution, is to be preferred.

The analysis of variance of the design is given. To illustrate the various procedures, a group of two experiments on different species of Eucalypts, each with five replications and ten treatments, one of which was common to both the experiments, is analysed.

(R. L. Chaddha)

GRIDGEMAN, N. T. (National Research Council, Ottawa) **9.5** (6.8)
Application of quantal-response theory to cross-comparison of taste-stimuli intensities—*In English*
Biometrics (1948) **14**, 548-557 (8 refs., 3 tables, 2 figs.)

This article concerns some taste testing experiments which were carried out using the method of paired comparisons. The purpose of these experiments was to illustrate the use of this technique when applied to a quantal response bioassay problem. In particular, various concentrations of sucrose were compared with various concentrations of sodium chloride, citric acid and quinine sulphate. These represent the four primary tastes.

The underlying model is based on Thurstone's " philog-gamma " hypothesis in which the use of the cumulative normal for the sigmoid is replaced by the logistic function. This is more in keeping with the assumption of the Fechner law that the strength of the sensory process is proportional to the logarithm of the stimulus. If C_i and C_j represent two different concentrations of a stimulus, then the model is :

$$Y = ln[P/(1-P)] = \alpha + \beta \log (C_i/C_j)$$

where P is the proportion of the time that C_i is judged stronger, say, than C_j. $E(\alpha) = 0$ and β is a parameter of sensitivity. The use of the Weber fraction in measuring acuity is also discussed and, in fact, is equal to

$$F = \text{antilog}\,[(ln,\ 3)/\beta] - 1.$$

For a given substance, $E(Y) = 0$ occurs at $C_i = C_j$. If, however, valid comparisons are being made between two qualitatively different stimuli, the concentration ratio at $Y = 0$ will be the taste-intensity ratio of the pure stimuli, weight for weight. This ratio is the main objective of this type of experiment. The appropriateness of the model can be studied by goodness-of-fit tests.

For the case where different substances are being compared, the model is then generalised to

$$Y = ln[P/(1-P)] = \alpha + \beta_i \log C_i + \beta_j \log C_j + \gamma_{ij} (\log C_i)(\log C_j)$$

where C_i and C_j are now concentrations of different substances. If the response to both substances is the same, then $\gamma = 0$, $\beta_i = \beta_j$ and the model would become the same as the original one.

The experimental results are tabulated and discussed. These include estimation of the relative intensity for the four substances which are then compared with earlier, less statistically rigorous, methods.

(J. E. Jackson)

MAURICE, RITA J. (University College, London) 9.5 (1.8)

Selection of the population with the largest mean when comparisons can be made only in pairs—*In English*

Biometrika (1958) **45**, 581-586 (7 refs., 4 tables, 1 fig.)

Where it is required to select the best (indicated by the largest mean) of a number of normal populations of known equal variance and these populations can be compared only in pairs, the usual procedure is to use the standard experimental design for incomplete blocks with two plots per block. A possible alternative procedure would be to use a procedure of successive elimination, following the procedure familiar in cup-ties and tournaments. The simplest case in which the latter procedure may be used is for the selection of the best one out of four and its performance in this situation is examined.

The effect of the blocks on the experimental results may be constant or subject to random variation. If the block effects are constant, the cup-tie procedure appears to attain a given probability of correct selection more cheaply than the incomplete blocks procedure. This held for a number of probability levels, but the general case could not be proved. Assuming that the loss resulting from an incorrect choice is proportional to the difference between the largest and the chosen mean, the minimax principle may be used to determine the sample sizes of both procedures. The expected loss of the cup-tie procedure is then less than that of the incomplete blocks procedure whatever the size of the difference between the largest and the other means.

If the block effect is a random variable, the incomplete blocks procedure achieves a given probability of a correct choice more cheaply than if the block effects are constant. The analysis of the cup-tie procedure is not affected by the change in the model. If the variance of the block effect (σ_b^2) is sufficiently small the incomplete blocks procedure will be cheaper than the cup-tie procedure. For the probability levels selected, σ_b^2 must be less than or equal to $0{\cdot}8\sigma^2$ (where σ^2 is the variance of the populations) for this to hold in all cases. When the minimax principle is used to determine the sample sizes the incomplete blocks procedure is always better if $\sigma_b^2 = 0{\cdot}5\sigma^2$, and always worse if $\sigma_b^2 = 2\sigma^2$. For $\sigma_b^2 = \sigma^2$ and $\sigma_b^2 = 1{\cdot}5\sigma^2$ which procedure results in the smaller expected loss depends upon the value of the difference between the largest and the other means. The comparisons are made assuming σ_b^2/σ^2 is known. In practice this ratio may have to be estimated and the incomplete blocks procedure will then compare less favourably with the cup-tie procedure.

(R. J. Maurice)

SCHEFFE, H. (University of California, Berkeley) 9.7 (9.3)

Experiments with mixtures—*In English*

J. R. Statist. Soc. B (1958) **20**, 344-360 (2 refs., 4 figs.)

The author discusses the design of experiments in which the response depends on the properties of the components present in a mixture and not on the amount of the mixture. The factor space for a q-component mixture with the experiment arranged to have $m+1$ equally spaced proportions $(0, 1/m, 2/m \ldots\ldots, 1)$ of the components, is a simplex lattice denoted by $\{q, m\}$. All the possible mixtures correspond to points on the lattice.

He considers a $\{q, n\}$ polynomial of degree n in q variables having the same number of coefficients, of the various products of the variables, as the number of points on a $\{q, n\}$ lattice. The author conjectures, proving it for the case of $n = 1$, 2 and 3, that the values of the polynomial can be assigned arbitrarily on the lattice and that its values on the simplex are then uniquely determined. Combinations of the coefficients of the polynomial may be interpreted in terms of responses to the mixtures; linear coefficients to pure components, quadratic coefficients to synergism between two components and so on. The author further shows that for $n = 1$, 2 and 3, the relation of the coefficients to the responses at points on the lattice is related to orthogonal polynomials of the responses and considers significance tests of the fit of the polynomials. Suggestions for the allocation of numbers of observation at each point are made from consideration of the variances of predicted responses based on the order of the polynomial selected. Modifications of simplex lattice designs are suggested when it is necessary to keep the properties of one or more of the components below certain limits.

In a mathematical appendix the following points are examined:

(i) The number of points on a $\{q, m\}$ lattice.
(ii) The number of coefficients in a $\{q, n\}$ polynomial.
(iii) The canonical form of a $\{q, n\}$ polynomial for $n = 1$, 2 and 3.
(iv) The calculation of certain orthogonal polynomials and the maxima of their squares.
(v) A statistic for testing several correlated differences between observed and predicted responses.
(vi) Estimation of the regression coefficients when $m > n$ and one observation is made at each point.

(N. W. Please)

SHAH, B. V. (University of Bombay) 9.1 (7.4)

A note on orthogonality of experimental designs—*In English*
Bull. Calcutta Statist. Ass. (1958) **8**, 73-80 (no refs.)

This paper adopts the definition of an orthogonal design as one which provides mutually uncorrelated estimates of any two estimable parameters belonging to two different groups.

For designs eliminating heterogeneity in one direction, the necessary and sufficient condition for orthogonality of estimable block contrasts and estimable treatment contrasts is found to be that n_{ij}/r_i should be constant for all i and n_{ij}/k_j constant for all j, n_{ij} being the number of times the ith treatment occurs in the jth block, r_i the replication of the ith treatment and k_j the number of plots in the jth block. Hence it is shown that orthogonality cannot be achieved in a connected incomplete block design.

A similar set of necessary and sufficient conditions for orthogonality is also derived for general designs involving three groups of parameters, which are used for elimination of heterogeneity in two directions.

A measure of non-orthogonality is defined in terms squares of extreme values of correlation coefficients between estimable block and treatment contrasts and its value computed for Balanced incomplete block and Group divisible partially balanced incomplete block designs.

(I. M. Chakravarti)

UMBERGER, E. P., GASS, G. H. & CURTIS, J. M. (Food & Drug Admin., U.S. Dept. Hlth., Washington) 9.7 (6.8)

Design of a biological assay method for the detection and estimation of estrogenic residues in the edible tissues of domestic animals treated with estrogens—*In English*
Endocrinology (1958) **63**, 806-815 (11 refs., 2 tables, 4 figs.)

Experiments are described to estimate the residual estrogenic activity arising from the administration of stilbestrol to the diet of laboratory mice. The logarithm of the ratio of uterine weight to final body weight was generally found to be linearly related to dose of stilbestrol.

The treated mice were those fed tissues taken from other animals which had been fed diets containing stilbestrol. The control mice were fed tissues taken from animals which had not been fed stilbestrol. Four experimental designs are described to compare the tissues of these two groups of mice.

A. *Two-curve method.* Stilbestrol was added to both control and treated tissue diets at concentrations of 0, 2, 4 and 6 parts per billion. Dosage response lines were obtained for both groups of mice. The amount of residual activity, M, in parts per billion of stilbestrol, was estimated from the horizontal distance between the two lines. This method could not be used if $M > 2$.

B. *Range-finder method.* If $M > 2$, control diets were used with 0, 3 and 6 parts per billion of stilbestrol added; five other groups were fed treated tissue diets diluted with 0, 50, 75, 87·5 and 93·75 per cent. of control tissue diet.

C. *Confirmatory dilution assay.* Based on the range-finder experiment, the treated tissue diet was diluted with control tissue diet so that the expected concentration of residual stilbestrol would be between 1 and 2 parts per billion. A new two-curve experiment was then conducted.

D. *Single point assays.* These are to be used only in case there is insufficient tissue to perform a two-curve experiment.

Preliminary experiments had indicated that rats required 3 to 4 times more tissue than mice in detecting given residual estrogenic activity. Statistical procedures are those of Bliss (*The Statistics of Bioassay*, 1952, Academic Press, Inc., New York).

(M. Carter)

ADAM, A. (Institut für Statistik der Universität Wien) **10.5** (7.0)
Entropy and variance—*In German*
Metrika (1958) **1**, 99-110 (10 refs.; 2 figs.)

The paper is intended to direct the interest of the statistician towards information theory. Some relations are shown between entropy, in the sense of Shannon and variance, as commonly used in statistics.

The author treats information theory based on axioms as formulated by Chintschin and Faddeyew ; he introduces entropy in Chintschin's definition and then gives Faddeyew's axiom 3′, which concerns the increase of entropy due to the addition of one more item (p_n is divided into $q_1 . p_n$ and $q_2 p_n$, $q_1 + q_2 = 1$). The author points out that an analogue relation holds true for the variance.

According to the author, the relationship between the calculus of variance and of entropy, respectively, would allow the formulation of several statistical methods, based on the calculus of variance, in an analogous way for cases where only the concept of entropy is applicable.

A second example is concerned with regression analysis. Again the formalisms in the two " languages " are given. It is shown that at least in special cases the coefficient of determination based on entropy may be better than a measure of contingency.

Finally, it is shown that for a given variance the normal distribution has maximal indeterminacy.

(H. Zemanek)

ALLING, D. W. (Cornell University, Ithaca, New York) **10.9** (10.2)
The after-history of pulmonary tuberculosis : a stochastic model—*In English*
Biometrics (1958) **14**, 527-547 (13 refs., 6 tables, 3 figs.)

Stochastic models of chronic illnesses have resulted from the realisation that the course of a disease exhibits regularity in groups of patients although it is unpredictable in individuals. Validity of use depends upon whether clinical events are genuinely random in nature. Application to chronic disease requires a clearly defined starting point in time (diagnosis, completion of treatment, relapse) and definition of mutually exclusive clinical states. If instantaneous transition between states is possible, the model may be made time continuous. Sometimes in chronic disease such instantaneous transition is precluded by the choice of clinical states. The model discussed in this paper contains two such states, arrested pulmonary tuberculosis and active pulmonary tuberculosis. The definition of arrested pulmonary tuberculosis requires that certain conditions be present during a six-month period before entrance into the state is possible ; otherwise, the disease continues to be classified as active. The conditional probability function for " patient has arrested disease at time t_2 " given that " patient had active disease at time t_1 ", may lack a derivative and may be discontinuous.

If inquiry as to state is made at specified times, a discrete stochastic process will result. The transition probability function will be discrete and questions of differentiability and continuity will not arise. The process may be stationary even if the continuous process was not. The present model is such a case. Discrete processes permit more freedom of definition and since clinical data are usually summarised at discrete intervals of time, no information is lost. To avoid too great complexity stationary Markov processes are a necessity in any but the simplest situations. Data studied have been truncated at a time period before death from other than tuberculosis to avoid non-stationarity. Accuracy of classification of cause of death is examined to ensure validity of truncation. A study is made of some follow-up data of pulmonary tuberculosis. The following six states were considered in the model (1) active tuberculosis that will remain active indefinitely, (2) arrested tuberculosis that will remain arrested indefinitely, (3) tuberculosis that has proved fatal, (4) arrested tuberculosis that is certain sometime to become active, (5) active tuberculosis that is certain sometime to become arrested, and (6) active tuberculosis that is certain sometime to be fatal without ever becoming arrested. Transition probabilities are assumed constant with some " impossible " transitions having zero probabilities. The problem is reduced to one of estimating six unknown probabilities. Maximum likelihood estimators are derived with suitable iterants suggested. The method is applied to the follow-up data and reasonably good agreement between observation and prediction is obtained. It is suggested that better agreement would result from a finer classification of " active " tuberculosis.

(M. E. Turner)

BARKER, J. S. F. (Department of Animal Husbandry, Sydney University) **10.9 (11.7)**

Simulation of genetic systems by automatic digital computers. Parts III and IV—*In English*

Aust. J. Biol. Sci. (1958) **11**, 603-612 & 613-625 (10 refs. 2 tables, 2 figs., & 22 refs., 3 tables, 3 figs.)

III. Selection between alleles at an autosomal locus.

This paper describes a Monte Carlo approach to analysis of the effect of selection on gene frequencies. An electronic digital computer, the SILLIAC at Sydney University, is used to simulate the selection processes operating in populations.

The programme simulates selection between two autosomal alleles, allowing for selection at four stages of the life cycle: (i) Zygote selection: ability to survive from fertilisation to sexual maturity; (ii) Reproductive selection: differential reproductive ability of different genotypes; (iii) Selection between gametes at meiosis in heterozygotes; (iv) Selection between gametes on their ability to take part in fertilisation. These, and the population size can be varied to investigate different genetic situations.

To test the programme, two experiments reported in the literature have been simulated: (i) Selection between ST and CH chromosomal arrangements in the third chromosome of *Drosophila pseudoobscura* [T. Dobzhansky and O. Pavlovsky, *Evolution* (1953) **7**, 198-210]; and (ii) Selection between glass and wild type in *Drosophila melanogaster* [D. J. Merrell and J. C. Underhill, *Genetics* (1956) **41**, 469-485].

Close agreement was obtained in (i): the importance of an adequate estimate of generation length in these comparisons is discussed. Agreement with the results of (ii) was not so close: possible reasons for this in terms of selective mating are discussed.

121

IV. Selection between alleles at a sex-linked locus.

A programme, simulating selection between two alleles at a sex-linked locus, and developed for SILLIAC introduces selection and chance effects at four stages of the life cycle.

Two experiments on competition between sex-linked alleles are simulated, selection between (i) the "sex ratio" and standard X-chromosomes of *Drosophila pseudoobscura* [B. Wallace, *Evolution* (1948) **2**, 189-217], and (ii) *yellow* and its wild-type allele in *D. melanogaster* [D. J. Merrell and J. C. Underhill, *Genetics* (1956) **41**, 469-485].

(i) Wallace's data includes estimated overall selective values, and selective values for various components of fitness from which specific selective values for the programme's four stages are calculated. Two sets of simulated populations, one using the overall, the other the specific selective values, gave similar results. Though completely unlike the experimental population, they were very similar to the theoretical curve of change in frequency of sex ratio. Difficulties of estimating selective values from studies of the components of fitness are discussed, the need for an overall measure of fitness being emphasised.

(ii) The simulated populations for competition between *yellow* and wild type agreed closely with the experimental. The importance of selective mating in this competition is considered.

The author concludes that it is possible to simulate selection between two alleles both at an autosomal and a sex-linked locus by using automatic digital computers.

(J. Gani)

BENNETT, J. H. (Genetic Department, University of Adelaide) **10.9 (0.0)**

The existence and stability of selectively balanced polymorphism at a sex-linked locus

Aust. J. Biol. Sci. (1958) **11**, 598-602 (9 refs., 2 tables)

In a large random mating population, and in the absence of selective differences, the genotypes at a completely sex-linked locus with two allelomorphs A and a have the equilibrium frequencies p^2AA, $2pqAa$, and q^2aa in the homogametic sex (female say), and pA and qa in the heterogametic sex (male), p and q being the population frequencies of the genes A and a respectively.

When there are selective differences in the genotypes, the corresponding equilibrium values of the gene ratios in the gametic outputs are given by $u_f = (1+t_A-2S_{aa})/(1+t_A-2t_AS_{AA})$ for the female and t_Au_f for the male, where t_A is the ratio of the selective values of A to a in the male, and S_{aa}, S_{AA} are the selective ratios of aa and AA to Aa in the female. Hence it is shown that the necessary and sufficient condition for the existence and stability of a balanced polymorphism at a sex-linked locus is that $1+t_A$ be greater than $2S_{aa}$ and $2t_AS_{AA}$. In particular when $t_A = 1$, the necessary and sufficient condition for the existence and stability of an equilibrium is that the heterozygote should be at a selective advantage over both homozygotes; but when t_A has any value other than unity, it is neither necessary nor sufficient for the heterozygote to be at a selective advantage with respect to both the homozygotes in order that a stable balanced polymorphism may exist.

These results are considered in relation to Wallace's [*Evolution* (1948) **2**, 189-217] data for populations of *Drosophila pseudoobscura* in which selection is acting on the sex-linked condition "sex-ratio". It is also verified that the occurrence of heterozygotes with a frequency greater than that of the homozygotes does not necessarily indicate that the heterozygote has a selective advantage over both homozygotes.

(B. R. Bhat)

BLACKWELL, D., BREIMAN, L. & THOMASIAN, A. J. (University of California, Berkeley) **10.5 (1.5)**

Proof of Shannon's transmission theorem for finite-state indecomposable channels—*In English*

Ann. Math. Statist. (1958) **29**, 1209-1220 (6 refs.)

The information source considered here has a finite number D of states, transitions between which are governed by an indecomposable D by D Markov matrix M. At the end of each transition one of an alphabet of A letters is emitted into an information channel. The channel has R states between which a transition then takes place, governed by one of AR by R Markov matrices $C(1), \ldots, C(A)$ selected by the input letter. A letter is emitted at the output of the channel, depending on the final state of this transition. The combination of source and channel is then a finite-state Markov process governed by a DR by DR matrix L. The channel is said to be *indecomposable* if the matrix L is indecomposable for every possible source. A definition of the capacity of such a channel is given. As a criterion for indecomposability it is shown that the channel is indecomposable if, and only if, every possible finite product of the channel matrices $C(1), \ldots, C(A)$ is an indecomposable matrix. It is shown that every channel with finite memory is indecomposable, so that the results of this paper include those of earlier work on finite-memory channels. The authors then prove a modification of McMillan's theorem, that there exists an entropy h for any Markov process governed by an indecomposable matrix, and that for almost all sequences of length N and probability p_n, $\lim_{N\to\infty}(-\log p_N)/N = h$. Shannon's theorem is proved in two parts, showing that through an indecomposable channel (*a*) it is possible to transmit at any rate less than capacity and (*b*) it is impossible to transmit at any greater rate. Finally, in another form of Shannon's theorem it is proved that an indecomposable channel of capacity H is " adequate " to transmit information produced by an ergodic source at rate h provided $h<H$, and that such a channel is not adequate if $h>H$; a precise definition of adequacy is given.

(C. W. Helstrom)

BRENY, H. (Université de Liège ; Centre d'Etudes de l'Energie Nucléaire, Belgium) **10.1 (10.2)**

An unfamiliar aspect of the stochastic theory of " dead "-time counters—*In French*

Ann. Soc. Sci. Bruxelles (1958) **72**, 81-96

This paper deals with the theory of " paralysis "-counters [W. Smith, *Proc. Camb. Phil. Soc.* (1957) **53**, 176] with a constant dead time τ and Poisson input. Denoting by t_0 an instant on which a particle has been registered, it is well known that the time-axis from t_0 onwards is divided into contiguous intervals, alternatively " covered " (with length τ), and " open " having a length which is a random variable with an exponential distribution. The situation one meets in going backwards along the time axis from t_0 is not symmetrical with the one going along the time axis from t_0 onwards, due to the way in which a counter behaves. According to the author it is because of this asymmetrical situation not being evident that the division in intervals as indicated does hold for the backward case. He then proves it to be correct : the proof proceeds by analysing the succession of events preceding t_0.

The paper ends with an application of these results to the problem of estimating the spectrum of the flight-time of particles.

(H. Breny)

FERRERI, C. (School of Statistics, Palermo) **10.2** (2.1)

On the generalisation of Vianelli's simultaneous conditional means—*In Italian*
Statistica (1958) **18**, 693-707 (7 refs., 1 table)

The author develops the statistical problem of conditional quasi-arithmetic means defined by Vianelli. A convenient application of this model to simultaneous power means is discussed in this paper. The opportunity of such an extension has also been considered in connection with other mean values.

Cases of this character occur in the analysis of statistical time series. For a set of n observed values in a time series x_{it} $(i = 1, 2, \ldots, n)$, each at different points of time t, the above-mentioned scheme is based on a *minimum* problem as follows

$$\Phi(X_i) = \sum_{i=1}^{n} \sum_{t=1}^{T} (x_{it} - X_i)^2 = \textit{minimum}$$

under the general condition

$$F(X_1, \ldots, X_n ; b) = 0$$

where the quasi-arithmetic means X_i, in the Vianelli sense, are unknown and b is known.

In order to determine X_i by the well-known Lagrange method, it is necessary to differentiate the linear relation

$$\Phi(X_i) + \lambda F(X_i ; b)$$

with respect to each of the $n+1$ unknowns $X_1, \ldots, X_n, \lambda$.

By analogy with the above-mentioned problem, the author suggests a general expression from which the quasi-arithmetic means follow as special case. This generalised scheme, based on conditional simultaneous power means, includes quasi-harmonic, quasi-geometric and quasi-quadratic means.

At the end of the paper the author gives some numerical examples principally concerning the Cobb-Douglas function in connection with the theory of marginal productivity and for which he uses the original statistical data.

(V. Amato)

GOODMAN, L. A. (University of Chicago) **10.3** (10.1)

Asymptotic distributions of " psi-squared " goodness of fit criteria for
mth order Markov Chains—*In English*
Ann. Math. Statist. (1958) **29**, 1123-1133 (16 refs.)

A probabilistic model for the random behaviour of an object which moves at discrete time points from one to another of a finite number s of states is said to be a Markov chain of order k (an integer $\geqq 0$) if the conditional (transition) probabilities of the object entering a certain state at time t, given the whereabouts of the object at the previous k instants of time, suffice (together with certain initial probabilities) to determine all other probabilities in the model. The case $k = 1$ is the usual Markov chain. If $k = 0$, the behaviour of the object at any time t is independent of its past.

Considerable effort has been spent during the past five years on problems of inference in Markov chains. These inference problems are of two types: (i) the problem of estimating the transition probabilities, and, (ii) the problem of testing the hypothesis that a given set of data has come from a Markov chain of a specified order. The paper being reviewed is one of a series of studies by the author into such problems. The present paper, however, is more of a consolidating digression from, rather than a continuation of the main trend of his research, since it is concerned with determining the validity of several conjectures made by I. J. Good.

In order to make inferences, the experimenter will observe the behaviour of the process in question, the data being therefore a sequence of integer random variables, $(X_1, X_2, \ldots, X_T)$ say, where the value of X_j denotes the state the object is in at time j. Let $u = (u_1, u_2, \ldots, u_m)$ be a set of integers and let f_u be the frequency of times the object passes through the states $u_1, u_2, \ldots, u_m$ consecutively during the first T units of time. These statistics are sufficient for most inference problems. Let f_u, n denote the expected value of f_u under the assumption that the process considered is a particular Markov chain of order n, and set $f^*_{u,n}$ equal to the expectation of f_u under the assumption that the process is the Markov chain of order n having for its transition probabilities their maximum-likelihood estimates computed from the observed data. For testing the hypothesis that a process is a Markov chain of order n on the basis of the f_u, the following chi-square-like goodness-of-fit criteria have been proposed;

$$\psi^2_{m,n} = \sum_u (f_u - f_{un})^2 / f_{u,n} \quad \text{and} \quad \psi^{*2}_{m,n} = \sum_u (f_u - f^*_{u,n})^2 / f^*_{u,n}.$$

For these tests to be practical, their limiting distribution $(T \to +\infty)$ should be known. Goodman proves the following conjecture of Good. Under the composite hypothesis that the process is a Markov chain of order n, the limiting distribution of $\psi^{*2}_{m,n}$ is that of $\sum_{i=1}^{m-n-l} i\, Yg_i$ where Yg_i is a chi-square random variable with $g_i = (i-1)^2 i^{m-i-1}$ degrees of freedom, all random variables being independent. A second conjecture by Good concerning the asymptotic distribution of $\psi^2_{m,n}$ is not proved but a modified version is stated and proved.

(Ronald Pyke)

GUPTA, H. C. (University of Delhi) 10.1 (10.2)

Diffusion by discrete movements—*In English*

Sankhyā (1958) **20**, 295-308 (5 refs., 2 figs.)

The problem of diffusion considered in this paper is that of the motion of a particle along a straight line with one of two possible velocities $u_1(>0)$ and $u_2(<0)$. It is governed by a specified initial distribution and a matrix of transitional probabilities. This can then be considered as a homogeneous Markov chain with two possible states. Let $\nu_1(\nu_2)$ be the time spent during the first n intervals in the state $u_1(u_2)$ when it ends up in the state $u_1(u_2)$. The first section of the paper is devoted to finding (i) the characteristic function, (ii) the probability distribution, (iii) the moments of the random variables ν_1 and ν_2. In the limiting case, when the discrete movements become continuous, the limiting values of the mean and mean-square are evaluated. The modifications necessary, in case immigration or emigration of particles is present, is touched upon.

In the next section differential equations governing continuous diffusion are obtained. It is a second-order partial differential equation which contains as particular cases, (i) the equation of heat conduction, (ii) the equation of motion of a sound wave, (iii) the equation of diffusion by continuous movements with a constant rate of emigration. The last section is devoted to the investigation of the probability distribution of $\nu_1(\nu_2)$ when one or more barriers (reflecting, partly reflecting and absorbing) are present. The method of images is employed to reduce it to the case when barriers are absent. Some of the results of this section are generalisations of certain results of Chandrasekhar.

(R. Ranga Rao)

HAIGHT, F. A. (Inst. Transportation & Traffic Eng., California) 10.4 (2.5)

Two queues in parallel—*In English*

Biometrika (1958) **45**, 401-410 (1 ref.)

Each new arrival to a system consisting of two queues A and B is assumed to join the shorter queue, or, if they are of the same length, one particular queue, say queue A. The arrivals are assumed to comprise a homogeneous Poisson process, and the service is also assumed to be Poissonnian (though not necessarily with the same parameters for the two queues). Differential-difference equations are set up and, assuming equilibrium, various relations are found between the state probabilities, the individual (marginal) queue-length distributions together with some of the lower moments of the conditional queue-length distributions. If the length of queue B is independent of time, the problem reduces to one previously studied by the same author in " Queueing with balking " [*Biometrika* (1957) **44**, 360-369].

Considerable simplification occurs when it is supposed that any member of a queue will instantaneously change queues whenever it appears advantageous to do so, i.e., when the other queue is the shorter. Thus the lengths of the queues can never differ by more than one unit. Explicit expressions for the equilibrium state probabilities are given ; each of the three possible configurations (equal queues, queue A longer, queue B longer) has (conditionally) a geometric distribution.

(C. L. Mallows)

HÁJEK, J. (Math. Inst., Czechoslovak Academy of Science, Prague) **10.5 (1.1)**

A property of $\mathcal{J}$-divergences of marginal probability distributions—*In English*

Czech. Math. J. (1958) **8**, 460-463 (5 refs.)

$\mathcal{J}$-divergence is an entropy-like measure introduced by H. Jeffreys to measure the difference of two probability distributions P and Q. $\mathcal{J}$ is the sum of relative entropies of P with respect to Q and of Q with respect to P. The paper is based on a rather general definition of $\mathcal{J}$-divergence, holding for probability measures on arbitrary fields of sets. The marginal distribution of a p-measure means the measure restricted to a subfield of the original domain of definition. So, for the marginal distribution, the $\mathcal{J}$-divergence is also well defined.

At first, certain properties of $\mathcal{J}$-divergence are listed. It is a symmetrical function of P and Q, it is always non-negative and equals zero if, and only if, both P and Q are identical. It is shown by a counter-example that $\mathcal{J}$-divergence cannot be used as a distance in spaces of measures because it does not possess the triangle property. Another property is a kind of monotony with respect to the Borel-field. That is, the $\mathcal{J}$-divergence of marginal distributions is always less or equal to the divergence of the original field. The equality holds if the marginal $\mathcal{J}$-divergence is infinite or, when the subfield is sufficient to distinguish between P and Q. Let us note that the divergence can be finite only if P and Q are equivalent measures.

The property announced in the title is the following: let us have an increasing sequence of Borel-fields and let F be the smallest Borel-field, containing every field of the sequence, then the $\mathcal{J}$-divergence of measures P and Q with respect to F is the limit of $\mathcal{J}$-divergences on the fields of the sequence. In the proof, the case of an infinite limit is first excluded as trivial; then it is shown that the measures P and Q are equivalent on F if the limit is finite. For the last step Doob's theorem on martingales is needed. By an analogous method similar results concerning continuity of relative entropies can be obtained. Also if the limit of relative entropies of P with respect to Q is finite, probability P is absolutely continuous with respect to Q on the field F.

The theorem, when applied to stochastic processes, gives the result that the $\mathcal{J}$-divergence between two distributions of a random function is the least upper bound of the divergences of finite-dimensional distributions. In this case the field F is the smallest Borel-field, generated by finite-dimensional cylinder-sets. It is shown that divergence of F equals that on a Borel-subfield generated by the values of the random function on a countable set of time-instants and the assertion easily follows. In an article in the next number of the same journal the author gives an application of this theory to the problem of deciding, with probability zero of error, between two Gaussian distributions of a random function, when one realisation of the function has been observed.

(P. Mandl)

HÁJEK, J. (Math. Inst., Czechoslovak Academy of Science, Prague) **10.0 (-.-)**

A property of normal distributions of any stochastic process—*In Russian*

Czech. Math. J. (1958) **8**, 610-618 (6 refs.)

In a previous paper in the same journal (*Czech. Math. J.*, **8**, 460-463) the author proved that $\mathcal{J}$-divergence (H. Jeffreys) of two probability distributions P and Q of a stochastic process is equal to the least upper bound of $\mathcal{J}$-divergences of finite-dimensional distributions of that process. Now, he uses this property to show that two normal distributions of a stochastic process can be only equivalent or singular. The distribution of a stochastic process is called normal when its finite-dimensional distributions are multivariate normal. Equivalency of P and Q means that a set has probability zero under P if, and only if, it is of Q-measure zero. P and Q are called singular if there exists a set which has P-measure one and Q-measure zero. In the case of singular distributions we can decide, with probability zero of error, between P and Q from one realisation of the process; so there are no statistical problems. The author's result can be restated in the form that, in any set of normal distributions of a stochastic process from which no two are singular, we have probability densities (Radon-Nikodym derivatives) of all measures with respect to any member of the class. The probability densities are of interest, for instance, when we want to form most powerful critical regions.

The proof is based on the following property of $\mathcal{J}$-divergence: to an arbitrary positive number e there exists a constant K such that we can decide from one observation, with probability of error less than e, between two arbitrary regular n-dimensional normal distributions of a random vector whose divergence is greater than K. When the divergence of two normal distributions is finite, they are equivalent. When it is infinite, because it is the least upper bound of finite-dimensional divergences and because of the property mentioned, a sequence of critical regions can be constructed, with probability of incorrect decision (both kinds of error) tending quickly enough to zero. From that sequence the region with probability zero of error is constructed.

The criterion of equivalence (finiteness of $\mathcal{J}$-divergence) is illustrated in two examples:

(1) Change of the scale parameter. In this case P and Q have the same mean-value function, but the covariance function of Q is that of P, multiplied by a positive constant C. The result is that when C is different from one, the measures are equivalent if, and only if, they degenerate in finite-dimensional distributions.

(2) Translation of a Gauss-Markoff process. In this case a necessary and sufficient condition of equivalency of two normal distributions, which differ only in their mean-value functions, is the finiteness of a Hellinger-integral.

(P. Mandl)

HANNAN, E. J. (Australian National University, Canberra) **10.6** (10.2)
The estimates of the spectral density after trend removal—*In English*
J. R. Statist. Soc. B (1958) **20**, 323-333 (7 refs., 2 tables)

Consider a discrete time-series which can be represented as the sum of (*a*) a multilinear regression on several known variables, with unknown coefficients, and (*b*) an independent stationary random process with an unknown spectrum. It is supposed that the regressor variables are such as to allow the (asymptotic) definition of autocovariance functions. Thus polynomial and trigonometric regressions are included. A further "orthogonality" condition is imposed on the regressor variables.

The problem is then to estimate the unknown spectrum when the regression coefficients are estimated by least squares—which is asymptotically equivalent to using the best linear unbiased estimates. This spectrum is needed to obtain the covariance matrix of the regression estimates.

The estimation of the regression coefficients introduces a first-order bias into the straightforward (unsmoothed) estimate of the spectrum from the residuals. This bias is evaluated and is shown to be a simple function of the spectra of the regression variables. It corresponds (asymptotically) to multiplying the straightforward estimate by a factor which can be calculated *a priori*. The sample autocovariances are also biased and these biases are similarly evaluated.

To obtain a consistent estimate of the spectrum it is necessary to smooth the straightforward estimate. This introduces a further bias which will be relatively negligible if the population autocorrelations decrease sufficiently rapidly. Both biases will be asymptotically negligibly relative to the standard error of the estimate. However, since the samples generally encountered in practice will be finite, it is thought worthwhile to investigate methods of eliminating them. Several suggestions are made, the simplest corresponds to multiplying the smoothed estimate by a factor (a smoothed version of that mentioned above) which again can be calculated *a priori*. This will not affect the order of the asymptotic variance of the estimate, but will reduce the order of the overall bias to that of the bias due to smoothing.

A numerical example is given; it is admitted to be favourable to the use of the proposed factor since the bias due to smoothing is relatively small.

(C. L. Mallows)

KLEIN, K. R. (Institute of Statistics, Oxford and Cowles Foundation, Yale) **10.2** (4.2)
The estimation of distributed lags—*In English*
Econometrica (1958) **26**, 553-565 (9 refs.)

The purpose of this paper is to advance the understanding of the lag scheme $y_t = \alpha \sum_{i=0}^{\infty} \lambda^i x_{t-i} + u_t$ proposed by L. M. Koyck [*Distributed Lags and Investment Analysis* (1954)]. In particular, the problem of estimating the parameters α and λ is considered. A derivative of the above equation is used which involves only three observable variables, y_t, y_{t-1} and x_t. Ordinary least squares procedures, however, leads to biased estimates of the coefficient parameters since the composite error term in the derived form of the model is not independent of y_{t-1} and also because the composite error terms are serially correlated.

Koyck's method of estimation, when the original series of error terms are not autocorrelated, is shown to be equivalent to a type of maximum likelihood estimate. An alternative computing method is suggested which enables the estimates of α and λ to be obtained in one step and not two. In addition a generalised statement of the problem, with zero autocorrelation in the original error terms, is given which enables the method of solution to be generalised to any number of variables.

The case of autocorrelation in the original series of error terms is then considered. When the autocorrelation parameter is equal to λ, it is shown that the least squares regression of y_t on x_t and y_{t-1} would give consistent and efficient estimates. In other cases it is necessary to transform the variables in order to obtain equations not involving autocorrelated errors. An equation system is developed which can be solved for α, λ and the autoregressive parameter by an iterative procedure.

The paper concludes with a general discussion of the suitability of the general model and its derivative for the investigation of particular problems in economics. It is suggested that with cross-section, as opposed to time-series data, the techniques for zero autocorrelation would be appropriate.

(D. J. Day)

KOŹNIEWSKA, I. (Planning & Statistics High School, Warsaw) **10.1** (1.5)
Ergodicity of non-homogeneous Markoff chains with two states—*In Polish*
Colloq. Math. (1958) **5**, 208-215 (6 refs.)

The scope of the note is to compare different notions of ergodicity for non-homogeneous Markoff chains with two states. The notions of ergodicity of a Markoff chain with r states, considered by Mrs Koźniewska, can be formulated in terms of transition probabilities $h_{ab}(m, n)$ from a state E_a at moment m to a state E_b at moment n as follows :

I. Ergodicity in Kolmogoroff's sense : for every $m = 0, 1, \ldots$ and any fixed a, b and c the difference between $h_{ab}(m, n)$ and $h_{cb}(m, n)$ tends to zero if n increases to infinity.

II. Weak ergodicity in Hajnal's sense : for any a, b, c the difference between $h_{ab}(0, n)$ and $h_{cb}(0, n)$ tends to zero with n increasing to infinity.

III. Strong ergodicity : for any a, $h_{ab}(0, n)$ tends with n increasing to infinity to a limit depending only on b.

IV. Strongest ergodicity : for every $m = 0, 1, \ldots$ and any a the probabilities $h_{ab}(m, n)$ tend with increasing n to a limit depending only on b.

Obviously strongest ergodicity implies the strong one and Kolmogoroff's ergodicity implies that of Hajnal. Mrs Koźniewska gives examples of two-state non-homogeneous Markoff chains which are ergodic (*a*) in all senses, (*b*) only in the sense of II and III, (*c*) only in the sense of II, (*d*) only in the sense of I and II, (*e*) in no sense. Then, using the necessary and sufficient conditions for a two-state Markoff chain to be ergodic in the four senses in question, formulated in terms of characteristic roots of matrices of transition probabilities from a state at moment n to a state at moment $n+1$, she gives conditions (f) for the ergodicity I to be equivalent to ergodicity II, and (g) for ergodicity III to be equivalent to ergodicity IV. Conditions are also given for a two-state non-homogeneous Markoff chain to be ergodic in the strong sense without being ergodic in the strongest sense ; to be ergodic in Kolmogoroff's sense without being ergodic in the strong sense ; to be ergodic in the weak sense without being ergodic in any other sense. Finally, the equivalence of the four kinds of ergodicity for two-state homogeneous Markoff chains is proved.

(S. Zubrzycki)

LEIPNIK, R. (Test Dept., US NOTS, China Lake, California) **10.7** (0.4)
Note on the characteristic function of a serial-correlation distribution—*In English*
Biometrika (1958) **45**, 559-561 (6 refs.)

The moments of a two-parameter distribution appearing in several studies of serial correlation have recently been discussed both by M. G. Kendall and by J. H. White.

The present author finds a Neumann-type series for the characteristic function of this distribution for a generalised parameter range. The derivation of the characteristic function is given and an expression is obtained for the pth moment about zero of the distribution. Agreement with Kendall and White is noted. The consequence of the standardised distribution to the normal, as one of the parameters tends to infinity, is demonstrated through the characteristic function.

The author remarks that " the technique of asymptotic differential equations employed may be useful in other such problems. It is familiar in physics under the name W.K.B. method."

(F. N. David)

LESLIE, P. H. & GOWER, J. G. (Bur. of Animal Pop. Oxford & Rothamstead Exp. Station) **10.9** (10.1)
The properties of a stochastic model for two competing species—*In English*
Biometrika (1958) **45**, 316-330 (11 refs., 5 tables, 4 figs.)

Conditions are found for the stationary values of the system in terms of the parameters of a discreet time, interacting, deterministic model. Inequalities which arise lead to three types of stationary state, (i) stable, in which both species persist, (ii) unstable, in which either one or the other species persists according to the initial state of the system and (iii) in which one species persists and the other disappears independently of the initial state.

The variety of stochastic models having the same deterministic equivalent may be considered to lie between two extremes : Model I in which the birth rate of each species is constant, and Model II in which the death rate of each species is constant. A stochastic process is constructed by attaching random normal deviates to the number of each species in existence at each stage of the deterministic development. As a consequence of this the outcome of a system with an unstable stationary state (ii) depends not only on the initial state but on the early development of the system. Using sixty replicates on each of a range of initial states, contour lines for the probability that a given species will survive (are estimated) for both Models I and II.

In the case of the stable stationary state (i) an analysis is carried out on the life histories of the two systems with the same initial states and deterministic model in which the birth rate of each species is constant (Model I). It is found that the fluctuation of the number of each species in existence about the stationary state follows a bivariate distribution which is approximately normal but that the variation about the stationary state is greater than that which may be predicted from theoretical considerations.

By using sixty replicates for each of a variety of initial states, an estimate is made of the probability that the stochastic outcome is the same as the outcome of a deterministic model with a type (iii) stationary state using Model I.

Life histories are obtained by reading into a computor the parameters of a deterministic model together with the initial conditions and a pair of random numbers. These generate the pseudo-random deviates required at each stage of the stochastic development.

(J. G. Saw)

LÉVY, P. (Ecôle Polytechnique, Paris) **10.1** (6.2)
On various classes of random functions—*In French*
J. Math. Pures Appl. (1959) **38**, 1-23 (8 refs.)

The author gives in this paper a complement to a theory of Laplacean random functions presented by him at the Third Berkeley Symposium and the more general study of linear correlation functions. The paper begins with work on random functions in time $\phi(t)$ which are, for all fixed (t), certain linear functions of values in $(0, t)$ of an additive random function.

The author represents these in the two forms of Stieltjes integrals ; then he shows that they are identical. The confines of this set of functions contain all the linear correlation functions.

The proofs use theorems of Frechet and Damois generalising the results of S. Bernstein. The paper contains eight theorems and the author points out the interest of this question for the problem of classification of random functions.

(D. Dugué)

MORAN, P. A. P. & WATTERSON, G. A. (Australian National University, Canberra) **10.9** (10.1)

The genetic effects of family structure in natural populations—*In English*

Aust. J. Biol. Sci. (1959) **12**, 1-15 (7 refs.)

This paper deals with the rate of approach to homozygosity of dioecious diploid populations whose generations do not overlap. The probability that a finite genetic population without mutations is not totally homozygous at the t-th generation ($t = 0, 1, 2, \ldots$) is asymptotically equal to $C\lambda^t$, where λ is called the rate of approach to homozygosity, and C is a constant.

For certain models of Wright [*Genetics* (1931) **16**, 97-159; *Proc. Nat. Acad. Sci., Wash.* (1933) **19**, 411-419], Moran [*Proc. Camb. Phil. Soc.* (1958) **54**, 60-71], and Watterson (unpublished data) with specific distributions of offspring per individual of the parent generation, the known values of λ are given. These results are generalised to cases with arbitrary offspring distributions, and with permanent marriage.

First, the family structure of populations of fixed size N is considered, with N_1 males and N_2 females where $N_1+N_2 = N$. The authors obtain the generating functions $P_1(z)$, $P_2(z)$ for offspring distributions of male and female parents respectively, in terms of the birth and death rates for each individual of specified age; the first two moments of these distributions prove to be of importance in subsequent theory. As examples, the generating functions applying to the previous population models are given.

The effect of family structure on the genetic behaviour of a population after a large number of generations is now studied; the population is assumed to be diploid, with a diallelic genetic factor. The autosomal case in a non-overlapping generation model with independent random mating between adults in each generation is considered first. It is shown how λ, the largest non-unit root of the stochastic matrix for the Markoff chain of genetic states, may be evaluated from a set of difference equations as approximately

$$\lambda = 1 - N^{-2}[N_1P_1''(1) + N_2P_2''(1)]/8.$$

This value remains unaffected when permanent marriage occurs, with $N_1 = N_2 = N/2$ and identical variances for both offspring distributions. Further results are stated for the sex-linked case.

Finally, though the general solution for overlapping generation models is not known, it is shown that λ is the same for a pair of models of Moran's type whether generations overlap or not. That this result does not hold in general has been shown by Watterson, using a second pair of equivalent models. The conditions required to extend results in non-overlapping to overlapping generation models require further investigation.

(J. Gani)

PARZEN, E. (Stanford University, California) **10.6** (10.2)

On asymptotically efficient consistent estimates of the spectral density function of a stationary time series—*In English*

J. R. Statist. Soc. B (1958) **20**, 303-322 (16 refs., 2 tables)

The author considers continuous and discrete time series which are stationary (in the wide sense) to order four, and constructs several general classes of estimates of their spectral density functions. Each of these estimates can be regarded as a weighted average over the periodogram, or as a weighted sum of the sample autocorrelations. There are two important types of estimate, called *exponential* and *algebraic*, corresponding to weight functions which satisfy certain fairly weak inequalities.

Considering the case where the period of observation tends to infinity, definitions are given according to which an estimate may be *pointwise consistent* or *boundedly consistent*; *uniformly consistent*; *uniformly-boundedly consistent*; *functionally-uniformly consistent*; and *integratedly consistent*—depending on the integrated mean square error (i.m.s.e.). Each of these estimates being of a certain *order*, it is possible to define the minimum integrated mean square error and the asymptotic efficiency of an estimate.

The highest order of consistency with which the spectrum can be estimated depends on the smoothness of the spectrum, as expressed by the rapidity of the decrease of the population autocorrelations as the lag tends to infinity. The results are very different according to whether the decrease is (*a*) exponential, or (*b*) algebraic. The asymptotic i.m.s.e.'s for estimates of the exponential type [for (*a*)] and of the algebraic type [for (*b*)] can be evaluated and give simple results. In the first case, the actual choice of a weight function is not critical, but in the second case it is very critical. The behaviour of the autocorrelation function need not be known at all completely in the first case, but in the second case it must be known fairly well before it is possible to construct an efficient estimate. The class of spectra which are consistently estimated (in an integral sense) by a given estimate can in each case be simply characterised.

Similar results are obtained for asymptotic mean square errors corresponding to the other definitions of consistency. The minimum i.m.s.e.'s are evaluated numerically for each of eight suggested estimates in the case of a simple Markoff process and for three sample sizes.

(C. L. Mallows)

PFANZAGL, J. (Institut für Statistik der Universität Wien) **10.0** (0.2)

On the decomposition of statistical series—*In German*
Metrika (1958) 1, 130-147 (11 refs., 2 tables, 7 figs.)

The paper deals with the analysis of statistical time series in components, starting from a type of series which can be represented by an additive composition of a smooth component (trend and business-cycles), a periodical component (season) and an irregular component.

The author stresses the fact that the methods for the analysis of time series are frequently based on the assumption that the values of the seasonal component at corresponding points (e.g. in January, February, etc.) of different periods (e.g. years) are either constant or change only in a special way, depending on the corresponding values of the smooth component. In the latter case the amplitude of the seasonal component is usually assumed to be proportional to the smooth component. There exist, however, series which do not comply with these assumptions. The method of A. Wald takes this fact into account by starting with a more general model: the periodical component is assumed to have a constant shape, its amplitude is assumed to change only gradually, but in an arbitrary way, i.e. independent of the development of the smooth component. The author varies this model in such a way that the amplitude may be subjected to sudden changes from one period to the other. In accordance with this, the season is defined as the product of a periodical function, and—differently from Wald—of a magnitude (amplitude) which can assume a different arbitrary value every year. Consequently, the values of the seasonal component have to be computed separately for every period. Under the usual assumptions on the behaviour of the irregular component this is possible if the smooth component can be approximated in every period by a linear function of time. By subtracting the seasonal component from the original series a composite series of the smooth plus the irregular component is obtained. These two components can be separated with the variate difference method. The accuracy of the method is examined on the basis of an artificial series and its applicability demonstrated by analysing the series of scarlet-fever deaths in the years 1949 to 1957.

The procedure developed by the author is a generalisation of Wald's method, as the assumption of an arbitrary change of the seasonal amplitude is less restrictive than the assumption that there is a continuous change in time. If the assumptions underlying Wald's model are fulfilled both methods lead to practically the same results.

The new method can be used especially for the analysis of series showing an unexpected change of seasonal fluctuations. This will mostly be the case with agricultural production, infectious diseases, etc. Finally, attention is drawn to the possibilities of applying such methods to the analysis of agricultural and industrial experiments, as in this field formally equivalent relations occur.

A correction of several misprints is subsequently given in this volume of *Metrika* (1958).

(O. Anderson, jun.)

PRÉKOPA, A. (Eötvös University, Budapest) **10.8** (1.4)

On secondary processes generated by a random point distribution of Poisson type—*In Hungarian*
Ann. Univ. Sci. Budapest—Sect. Math. (1958) 153-170 (13 refs.)

In this paper an exact mathematical model of secondary stochastic processes and a condition under which the secondary process will be of " Poisson's type " is given.

Let T be an abstract space and $\mathcal{S}_T$ a σ-algebra of certain subsets of T. Suppose that in T a random point distribution of Poisson type is given, i.e. a random selection of a finite number of points such that if $\xi(A)$ denotes the number of random points lying in $A \epsilon \mathcal{S}_T$ then

$$P[\xi(A) = k] = (1/k!)[\lambda(A)]^k e^{-\lambda(A)},$$

where $\lambda(A)$ is a finite measure on the elements of $\mathcal{S}_T$. To every random point in T there corresponds a secondary " happening " i.e. a random selection of exactly one element of another abstract space Y in which a σ-algebra $\mathcal{S}y$ is given. If the random points are $t_1, t_2, \ldots, t_n$ and the corresponding " happenings " are $y_1, y_2, \ldots, y_n$ respectively, then a sample element of the whole secondary process is $(t_1, y_2), \ldots (t_n, y_n)$. Thus the secondary process is not else than a random point distribution in the product space $T \times Y$. The author proves that under a suitable condition on the space T and under the condition of independence of secondary happenings corresponding to different random points, the random point distribution in $T \times Y$ is also of Poisson type. If $\mu(C, t)$ is the distribution of the random happening y having the starting point $t \epsilon T$ and $A \epsilon \mathcal{S}_T$ then the parameter of the probability distribution of random points lying in $A \times C$ is

$$\int_C \mu(A, t)\lambda(dt).$$

The parameter corresponding to other sets of $\mathcal{S}_T \times \mathcal{S}_y$ can be obtained by extension of the preceding measure. Applications are mentioned, e.g. to a telephone problem, to the transparentness of a film darked by emulsion spheres and to wandering random points in a Euclidean space. The latter contains as a special case the result of Doob. (*Stochastic processes*, New York-London, 1953, pp. 404-407.)

(P. Rèvész)

RAMAKRISHNA, B. S. (Indian Institute of Science, Bangalore) **10.5** (–.–)
Information Theory and some of its applications—*In English*
Current Science (1958) **27**, 376-385 (11 refs., 1 fig.)

This is an expository paper dealing with Information Theory, as was developed by Claude Shannon, and some of its applications to the transmission of messages over channels with and without noise together with problems of language and human communication. The model of a general communication system is as that adopted by Shannon, viz. " the information source " from which message passes to the transmitter which releases the transmitted signal with " noise " added, and the received signal passes through the receiver which gives rise to message again. This message comes to the " destination ". Various properties of the quantity of information—$\Sigma p_1 \log_2 p_1$, where p_1, p_2 ... are the probabilities of n symbols, are enunciated to justify this particular definition of information. The capacity of a noisy channel is defined in the usual manner by maximising " the amount of information sent less the uncertainty that remains of the transmitted message after the message is received at the destination ". To use the channel to transmit information at any rate near to its capacity, one should assign probability $p_{(i)}$ to the input symbols such that $\Sigma p_{(i,j)} \log p_{(i,j)}/p_{(i)} \cdot p_{(j)}$ is equal to the required rate where $p_{(i,j)} = p(j \mid i)$. $p_{(i)}$, $p_{(j)} = \Sigma p_{(i,j)}$ and $p(i \mid j)$ is the channel probability: i denoting the input symbol and j the output symbol.

Another important branch to which the theory of information can be applied is that of language and human communication. Information theory has been developed without attention to the meaning of the messages. But when human communication is considered it is remarked that the semantic problems involved cannot be ignored. But nevertheless the problem of semantic information is not entirely separable from the " information of the telecommunication problem ", sometimes called the " selective information ". For if we have the choice between two languages in which to transmit the same semantic information, it is suggested that language which requires the smaller number of bits of selective information should be preferred. By adopting this method the author states that he has succeeded in studying the relative efficiencies of different Indian languages for communication of the same semantic content. A note on the relationship between Informational and Theromodynamical entropies is provided at the end.

(K. R. Parthasarathy)

SMITH, W. L. (North Carolina State College, late Statist. Lab., Cambridge) **10.0** (1.6)
Renewal theory and its applications—*In English*
J. R. Statist. Soc. B (1958) **20**, 243-284 (98 refs.)

This is a long review paper with nearly one hundred references: it is followed by 18 pages of discussion for which a separate abstract is given. The topic reviewed is that of " renewal theory " (alias the theory of recurrent events). A succession of " lifetimes " X_1, ... X_n ... is envisaged, each an independent sample from the same law, and interest is focused on the distribution of the number N of such events occurring in a time t (measured from the start of the first lifetime). The review is divided into four chapters. The first treats the fundamental theory based on the " renewal function " $H(t) = \mathscr{E}(N_t)$ and here the emphasis is on rigorous proof and generality (which is found to exist chiefly in the limit as $t \to \infty$). The second chapter discusses " regenerative stochastic processes " on the basis of Smith's result that the embedded renewal process provides a proof of the existence of an equilibrium limiting distribution for the regenerative process. This is also done for semi-Markoff processes. The chapter closes with a brief treatment of cumulative processes (partial sums of lifetimes), their normal and degenerate limits.

The third chapter treats four recent problems of a more particular and compound nature. The first is one of Takacs' (arising from Geiger counter theory) where the expectation of an infinite sum of a function of two independent renewal processes is required together with its auto correlation function. Asymptotic results are given together with a discussion of the corresponding spectral density function. The second problem is of the expectation $P(t)$ of an infinite product related to Takacs' sum (which has a death and immigration process application) and an integral equation is found for this. A differential equation is found for the corresponding Laplace Transform and its numerical solution is discussed together with explicit solution for the case of an underlying Poisson process. A series expansion for the asymptotic value of $P(t)$ is found in another special case. The third problem is that of the application of " paralysis " and " guarantee " censors to a renewal process and a critical catalogue of some thirty relevant papers is made. A variety of more or less complete solutions is found for various special underlying distributions. The fourth problem is Harris's treatment of branching renewal processes where, at the end of each lifetime, two new lives start and an integral equation for the mean population size at time t together with an asymptotic expression is given for this. The fourth chapter covers: Chang's extension of renewal theory to vector X s; Cox and Smith's treatment of the case where the $\{X_t\}$ have different p.d.f.s.; Wold and Doub's work on the case where the p.d.f. of X_t depends, in a Markovian fashion, on the previous values of X.

(D. E. Barton)

SMITH, W. L. (Report of a discussion on paper by) Renewal Theory and its Ramifications—*In English* *J. R. Statist. Soc.* B (1958) **20**, 284-302

10.0 (1.6)

The discussion on this review paper consists of nine verbal contributions, two written contributions received after the meeting and the author's detailed reply. Those who took part, in order of speaking, were :—M. S. Bartlett, D. R. Cox, J. M. Hammersley, J. G. Skellam, L. R. Shenton, A. Mercer, R. Syski and D. G. Kendall. The written contributions were from L. Takács and S. Karlin.

Professor Bartlett brought forward a " portmanteau " formula concerning renewal theory and also referred to the forthcoming publication of work relevant to m-dependant renewal process. Dr Cox discussed three generalisations including multidimensional renewal theory and a practical application involving renewal theory and the theory of order statistics. Mr Hammersley's contribution was extensive and dealt in some detail with the method of treating an integral equation (section 3.9 of the paper) and of the approach of renewal and associated processes through integral equations for cumulative distribution functions. Mr Skellam discussed a class of immigration-death processes where the arrivals are at random. Dr Shenton commented upon a renewal process which arose in polymer chemistry and Mr Mercer a practical application to conveyor belting damage. In this latter case explicit formulae are not available and attention must be turned to non-parametric results.

Mr Syski pointed out that the concept of regeneration points was introduced by Palm in 1943 [*Ericsson Technics* **44**, 1] and gave certain additional references to work in connection with equations for infinite sums. (See section 3.1 of the paper.) Mr D. G. Kendall's contribution noted that there was a good deal of common ground between renewal theory and the theory of Markov chains. He had carried out work on the rate-of-convergence problem associated with a limit theorem due to Kolmogorov and was interested to learn whether anything had been done in this connection with the analogous theorem of Blackwell. He also pointed out that Palm's term for what is now known as " regeneration point " was " point of equilibrium ".

Dr Takács and Professor Karlin provided written contributions to the discussion. Dr Takács showed how formulae (1.1) of the paper could be obtained without recourse to the Strong Law of large numbers and emphasised the key importance of the theorem (1.3) in the theory of stochastic processes. He gave an application of renewal theory to telephone traffic and a generalisation of renewal theory in connection with industrial processes.

Professor Karlin's written contribution drew attention to a further source of stochastic models whose analysis must rely upon renewal theory—inventory and production models.

The author's detailed reply summarised the discussion and offered comments further upon points raised by the participants.

(Wm. R. Buckland)

TAKÁCS, L. (Eötvös University, Budapest)

10.4 (–.–)

On a combined waiting-time and loss problem concerning telephone traffic—*In Hungarian*

Ann. Univ. Sci. Budapest—Sect. Math. (1958) 73-82 (8 refs.)

Suppose that at a telephone centre calls are arriving in the instants $\tau_1, \tau_2, \ldots, \tau_n, \ldots$. There are m full-availability channels and there is a waiting facility. If all channels are busy and the number of waiting calls is less than w (w is the maximum number of waiting calls), then each new call joins the waiting line and waits until a channel is free. If all m channels are busy and w calls are already waiting then any new call is lost.

We say that the system is in state $E_k (k = 0, 1, 2, \ldots m, \ldots, m+w)$ if the number of calls being served and waiting is k. Define the random variable $\eta(t)$:

$$\eta(t) = k \ (k = 0, 1, \ldots, m, m+1, \ldots, m+w)$$

if at the instant t the system is in state E_k. Put $\eta(\tau_n - 0) = \eta_n$. The problem is to determine the limiting distributions

$$\lim_{n\to\infty} P\{\eta_n = k\} = P_k, \qquad \lim_{t\to\infty} P\{\eta(t) = k\} = P_k^*.$$

The author determines these limiting distributions if the following assumptions are made :

(*a*) The instants τ_n form a sequence of recurrent events, i.e. the random variables $\tau_{n+1} - \tau_n$ ($n = 1, 2, \ldots$) are identically distributed and independent. Their common distribution function has a finite first moment.

(*b*) Denote by χ_n the holding time of the call arriving at the instant τ_n. Suppose that χ_n are independent random variables which are independent also of the process $\{\tau_n\}$. We suppose that the conditional distribution function $P\{\chi_n \leqq x \mid \eta_n = k\}$ ($k = 0, 1, \ldots, m+w-1$) is an exponential one with the same parameter μ for each k.

The basis of the determination of the limiting distribution P_k is that the sequence η_n forms a Markoff chain. Further, if the distribution function of the random variables $\tau_{n+1} - \tau_n$ is not a lattice distribution then the limiting distribution P_k^* is also determined. It is shown in the paper that P_k and P_k^* ($k = 0, 1, \ldots, m+w$) are independent of the initial state.

Finally the distribution function of the waiting time under the condition that a call is not lost is determined.

(J. Mogyoródi)

TAKÁCS, L. (Eötvös University, Budapest) **10.8** (1.4)

On the probability distribution of the measure of the union of random sets in a Euclidean space—*In Hungarian*

Ann. Univ. Sci. Budapest—Sect. Math. (1958) 89-95 (8 refs.)

Consider certain random points $\{Pi\}$ distributed in the Euclidean space of m-dimensions. To each random point let there be a corresponding Borel-measurable random set Hi in this space. Let this random set depend on a random parameter α_i. It is supposed that these random parameters are identically distributed independent random variables. Further suppose that the indicator functions of the random sets can be expressed by the same Baire-function $f(r, a)$, where $r = (r_1, r_2, \ldots, r_m)$ denotes the points of the space and a is a positive real number. Namely, if $\alpha_i = a_i$ and r_i denotes the radius-vector of Pi, then we have

$$f(r - r_i, a_i) = \begin{cases} 1 \text{ if } r \in Hi \\ 0 \text{ if } r \notin Hi. \end{cases}$$

Let $\zeta(x) = 1$, if $x = (x_1, x_2, \ldots, x_m)$ is a point of at least one random set and $\zeta(x) = 0$ otherwise. If H is a fixed set with finite positive Lebesgue-measure in the Euclidean space of m-dimensions, then the expression

$$\Theta(H) = \int_H \zeta(x)dx$$

gives the measure of those points of H, which are covered by at least one random set. $\Theta(H)$ generally is a random variable and the problem is to determine the distribution of $\Theta(H)$ if different assumptions are made about the distribution of the random points. The problem is investigated in the following four cases :

(a) The random points are distributed only on the set H according to a Poisson process with density λ, i.e. if $\nu(H)$ denotes the number of random points in H, then

$$\text{Prob}\{\nu(H) = n\} = (1/n!)\{\lambda\mu(H)\}^n \exp[-\lambda\mu(H)]$$

where $\mu(H)$ is the Lebesgue-measure of the set H.

(b) The random points are distributed on the whole space R_m according to a Poisson process with density λ.

(c) The random points are distributed only on H. It is supposed that $\nu(H) \equiv n$ and the random points are distributed independently of each other and uniformly on H.

(d) The random points are distributed only on H. It is supposed that $\nu(H)$ is an arbitrary random variable which assumes integer values and under the condition $\nu(H) = n$ the n random points are distributed independently of each other and uniformly on H.

In the four cases mentioned above the author determines the moments of $\Theta(H)$:

$$M\{[\Theta(H)]^s\} = \int_H \cdots \int_H \text{Prob}\{\zeta(x_1) = 1, \ldots, \zeta(x_s) = 1\} dx_1 dx_2 \ldots dx_s$$

where

$$\text{Prob}\{\zeta(x_1) = 1, \ldots, \zeta(x_s) = 1\} = \sum_{k=0}^{s} (-1)^k \sum_{1 \leq j_1 < j_2 < \ldots < j_s \leq s} p(x_{j_1}, x_{j_2}, \ldots, x_{j_s})$$

Here the probabilities $p(x_1, x_2, \ldots, x_k)$ are expressed by the indicator functions of random sets and the distribution function of the random parameters. The problem is thus solved, since the distribution function of $\Theta(H)$ is uniquely determined by its moments. (J. Mogyoródi)

TAKÁCS, L. (Eötvös University, Budapest) **10.3** (10.9)

Probabilistic treatment of meteoropathological phenomena—*In Hungarian*

Mag. Tud. Akad. Kut. Mat. Intézet. Közleményei (1958) 129-140 (7 refs., 3 tables)

Let us suppose that certain climatic phenomena (e.g. meteorological frontal passages) occur at the instants t_1, t_2, ..., t_n, ... and certain biological phenomena occur at the instants $u_1, u_2, \ldots, u_n, \ldots$. It is supposed that every meteorological front gives rise to an effect extended over an interval with length α (the instants t_n are the centres of the corresponding intervals). One can observe the instants of the meteorological fronts and of the biological phenomena in the time interval $(0, T)$ and the problem is to decide in what degree the biological phenomena can be considered as an effect of the meteorological fronts.

The mathematical model of the problem is the following :

Consider the case when the time differences $t_{n+1} - t_n$ are identically distributed, independent, positive random variables with distribution function $F(x)$. We suppose that the climatic phenomena have no effect. In this case it is reasonable to suppose that the sequence of the biological phenomena $\{u_n\}$ forms a Poisson process with density μ and $\{u_n\}$ is independent of the process $\{t_n\}$.

Denote by ν_T the number of those biological phenomena in the time interval $(0, T)$ which fall into the intervals corresponding to the meteorological phenomena. Further denote by ν_T^* the actually observed number of the biological phenomena. Then the probability $P\{\nu_T \leqq \nu_T^*\}$ may be considered as a measure of the effect of the meteorological phenomena.

The author proves that if T is sufficiently large then

$$P\{\nu_T \leqq \nu_T^*\} \sim \phi(\nu_T^* - MT/DT^{\frac{1}{2}})$$

where $\phi(x) = 1/(2\pi)^{\frac{1}{2}} \int_{-\infty}^{x} e^{-t^2/2} dt$ and $\lim_{T \to \infty} M\{\nu_T\}/T = M$

and $$D^2 = \lim_{T \to \infty} D^2\{\nu_T\}/T.$$

$M\{\nu_T\}$ and $D^2\{\nu_T\}$ may be determined with the aid of the distribution function $F(x)$.

Finally the problem is treated whether the biological phenomena can be considered as an effect of some climatical phenomena acting continuously with intensity $\lambda(t)$, $(0 \leqslant t \leqslant T)$.

(J. Mogyoródi)

TANNER, J. C. (Road Research Laboratory, D.S.I.R., England) **10.4 (2.5)**
A simplified model for delays in overtaking on a two-lane road—*In English*
J. R. Statist. Soc. B. (1958) **20**, 408-414 (2 refs., 1 table)

The model considered is of vehicles travelling in both directions along a long straight two-lane road. The vehicles travel at constant speeds of v in one direction and V in the other and are assumed to be spaced at random in their streams; the density of vehicles in the two streams being q and Q. The problem investigated is the delay experienced by a single additional vehicle during uninterrupted travel in the first direction at a constant speed u ($u \leqq v$), but which may only pass a v-vehicle according to the following rules:

(i) If the distance between the v-vehicle and the next approaching V-vehicle is d or more then it passes without slowing down.

(ii) If this distance is less than d it slows down instantaneously to v and waits until there is a distance D to the next approaching V-vehicle before overtaking.

(iii) It overtakes v-vehicles one at a time.

A function is obtained for the average speed u achieved by the u-vehicle in the course of a long journey along the road. This function is tabulated for a few selected values of the parameters. For specified speeds of the vehicles, expressions are also obtained for d and D—the minimum gaps in the opposing traffic which will allow passing to be safely performed.

The author points out that the model does not allow for the bunching which occurs in practice, owing to the variations in the speeds of vehicles in one direction, but suggests that it gives a fair approximation for roads with fairly light traffic. He mentions possible extensions to the model, to allow for bunching, by considering each stream of vehicles to be fed into the road as the output from a fixed—service—time-queueing process with random input. He points out that, though this type of bunching is different from that occurring in practice, the theoretical analysis of more realistic bunching is beyond the scope of present techniques.

(N. W. Please)

URBANIK, K. (Math. Inst., Polish Academy of Science, Warsaw) **10.1 (–.–)**
Remarks on invariant functions in Markoff processes—*In Polish*
Colloq. Math. (1958) **5**, 223-230 (3 refs.)

The note referred to is one of a series of Urbanik's papers devoted to a study of limiting properties of Markoff processes.

There are considered homogeneous Markoff processes whose sample functions are step functions of continuous time t assuming their values in a finite or denumerable set. Transition probabilities are assumed to depend continuously on time. A real valued function f, defined on the sample space of the process, is called invariant if it is measurable with respect to the class of sample function sets invariant under the shift transformations. The purpose of the note is to examine the power of a set of values which can be essentially assumed by an invariant function. Let us call the function $Q(x)$, defined for values x of the process as a limit of a probability that the process takes on x as its value at moment t if t tends to infinity, the limiting distribution of the process. Two theorems are proved. The first states that if the sum of all limits $Q(x)$ extended over all values of the process equals unity, then the invariant functions of the process assume essentially at most as many values as there are elements in the set of values of the process. In other words, if c is the power of the set of values of the process, then for each invariant function there exists a set of power c of its values such that it assumes a value from that set with probability one.

The second theorem states that for each sequence q_1, q_2, ... of non-negative real numbers with sum q less than unity, there exists a Markoff process with properties described above for which the limits $Q(x)$ assume exactly all values of the sequence q_1, q_2, ... Thus their sum, being equal to q, is less than unity; moreover there exists an invariant function of this process which assumes essentially non-denumerably many values, i.e. for each denumerable set of values of that function the probability that it assumes a value from this set is less than unity.

(S. Zubrzycki)

WETTE, R. (Zool. Inst., Biomath Dept., University of Heidelberg) **10.9** (2.5)

On the biomathematical basis of the logarithmic series distribution of elements of taxonomic units in the natural system—*In German*

Biom. Zeit. (1959) **1**, 44-50 (7 refs.)

Following a short discussion of Kendall's and Skellam's models for the generation of the empirically observed logarithmic series distribution of species per genus by stochastic processes in species evolution, a different model is suggested: the allocation of species to taxonomic genera by taxonomists is a stochastic process continuous in time with the following properties:

(1) The probability of allocating an additional species to an existing taxonomic genus is proportional to the size of the latter (probability of average likeliness),

(2) The probability of withdrawal of one species from a genus is proportional to the size of the latter (number of possible choices), with the exception that

(3) Unspecific genera (of size unity) tend to persist or when being abolished change their name with finite probability.

This is a modified Feller-Arley "birth-and-death" process, the modification consisting of a reflecting barrier between zero and unity. The author avoids integration of the derived partial differential-functional equation for the probability generating function by considering only the limiting stationary distribution. This stationary distribution is a logarithmic series with a zero term, the latter being censored in the empirical distribution.

(R. Wette)

ZÍTEK, F. (Math. Inst., Czechoslovak Academy of Sciences, Prague) **10.0** (1.6)

Random Functions of an interval—*In French*

Czech. Math. J. (1958) **8**, 583-609

The author considers random functions of an interval and their Burkill integral. In the paper it is supposed that random variables associated with any two adjacent intervals (i.e. two disjunct intervals the union of which is an interval) are stochastically independent. A random function of an interval is called additive if it associates with the union of two adjacent intervals the sum of random variables associated with separate intervals. The author introduces a notion of continuity in the following sense: the random variables associated with a sequence of intervals the length of which converges to zero converge to zero in probability. The Burkill integral is defined as follows: we divide a given interval into a finite number of subintervals, add the corresponding random variables and let the maximum length of subintervals tend to zero. If the corresponding sums tend in distribution or in probability to a certain limit, we say the random function has a (BB)-integral or a (pB)-integral, respectively, over the interval considered. The function is called integrable, if it has an integral over all subintervals of the given interval. Then the indefinite Burkill integral is an additive random function. Subsequently, the author gives some criteria for the existence of Burkill integrals of random functions, among them a criterion for the case of continuous random functions which is based on the theory of limit laws of probability.

The notion of (BB)-integral is used to make more clear some aspects of the theory of stochastic differential equations in the interpretation of P. Lévy. The author starts with the fact that a differential equation determines principally the distribution law of increments of a random function to be obtained. These increments can be written in the form of a random function of interval. By integration of this function we get an additive random function which is the solution of the differential equation considered.

(M. Sova)

Remark of the Regional Editor:

The results of the paper are partly overlapping with recent results of: A. Prékopa (1957), On stochastic set functions, III [*Acta Math. Acad. Scient. Hung.* **8**, 375-400].

(A. Rényi)

BARNARD, G. A. (Imperial College, London) 11.9 (1.0)
Thomas Bayes.—A Biographical Note—*In English*
Biometrika (1958), **45**, 293-295

Biographical information regarding Thomas Bayes, F.R.S. (1702-1761), the father of inverse probability, is very scanty. The author has collected together such information as there is and presented it as a short three-page preface to the reproduction of Bayes' essay presented to the Royal Society by Richard Price in 1763. Bayes, the son of a Presbyterian minister, was educated privately and became himself a Presbyterian minister in Tunbridge Wells. It is not known when he went to Tunbridge Wells, he was there in 1731, and he died there in 1761, having retired from his ministry in 1752. He was buried with his parents and brothers and sisters in Bunhill Fields.

(F. N. David)

FOSTER, F. G. (London School of Economics) 11.1 (5.8)
Upper percentage points of the generalised Beta Distribution, III—*In English*
Biometrika (1958) **45**, 492-503 (3 refs., 1 table)

The problem discussed in this and two previous papers by the author is that of testing a null hypothesis for k-variate normal populations by considering two independent estimates A and B, of what, on a null hypothesis, is the same dispersion matrix. The roots λ of the determinantal equation of $B-\lambda A$ are invariant under all linear transformations of the k variates, and it has been argued that some functions of the roots might be a useful criterion for testing the null hypothesis.

The author has been concerned with the tabulation of the greatest root λ, or, what is shown to be equivalent, of the greatest root θ of the determinantal equation of the difference between the sums of products matrices of B and A. θ, a function of λ and the degrees of freedom of A and B, increases monotonically with λ, and all θ lie between 0 and 1. The distribution function of the greatest θ depends on k, p and q (where p and q are respectively linear functions of the degrees of freedom of A and B) and is given the name of " generalised Beta Distribution " since it reduces to the Beta Distribution when k is 1.

The 80, 85, 90, 95 and 99 per cent. points were tabulated with k having value 2 in the first paper (published jointly with D. H. Rees). The value of k in the second was 3, and 4 in the tabulations of the present paper. The degree of freedom of A and B take values 5-195, 4-11 respectively.

Computations were carried out on the English Electric Company's DEUCE, as they were for the second paper, to which reference is made for details of the programme and interpolation requirements. Modifications of the computing procedure for k equal to 4 are given in this present paper. The existing programme could now be used to extend any of these three tables to higher values of the degree of freedom and hence of p and q. However, it is probable that this method would cease to be useful when k exceeds, or even equals 5, and a new theoretical approach appears necessary.

(D. Allibone)

HOTELLING, H. (University of North Carolina) **11.9** (1.0)

The statistical method and the philosophy of science—*In English* *Amer. Statistician* (1958) **12**, 9-15.

This is a philosophical discussion of probability and statistics which was presented as a presidential address to the University of North Carolina chapter of Sigma Xi. Accordingly it is directed at an audience of general scientific (rather than statistical) background. The aim of the paper is to indicate the impact that recent statistical theories are likely to have on the philosophical disciplines of metaphysics, epistemology and inductive logic.

The distinction between subjective and objective ideas of probability is discussed with particular reference to the theories of Keynes and Carnap. The axiomatic approach of Kolmogoroff is recalled. The author suggests that it may be appropriate to take "probability" itself as an undefined term, since every formal theory must necessarily contain certain undefined concepts. It is pointed out that the representation of ignorance by uniform probability distribution over a finite number of points may sometimes have justification via ergodic theory; as an example, in the shuffling of a pack of cards a uniform distribution of permutations is approached as the shuffling time is increased.

The philosophy of the usual techniques of point and interval estimation and of testing hypotheses is discussed. The impact of Fisher's *Design of Experiments* and of recent developments in sequential experimentation are recalled. Finally it is pointed out that there is a parallel between the pragmatic philosophy of William James and the theory of decision functions: One's stated objective is not to discover a true value, but rather to minimise a loss.

(R. J. Buehler)

KRISHNA MURTHY, E. V. (University of Madras) **11.6** (0.1)

LILAVATI—A new analogue computer for solving linear simultaneous equations and related problems. Part II. Design of Model II and its application to the solution of secular equations—*In English* *Proc. Ind. Acad. Sci.* (1958) **48**, 269-283 (3 refs., 8 figs.)

The general principles involved in the design of LILAVATI were discussed in a previous paper (in the same volume of the same journal) by the author jointly with G. N. Ramachandran. This second paper deals with an improved model of the first computer. The new design employs potentiometers instead of the meters in the Model I and so has the following advantages over it: (*a*) considerable economy is possible in material and in number of operations (*b*) the circuit can easily be modified to solve secular equations (*c*) operations can be made with a.c. and d.c. supply without changing the circuit appreciably.

Another advantage is that in this model, values of currents need not be read at intermediate stages but can be measured when the iterative process has converged. The technique of measurement is based on a null setting so that it is capable of great accuracy in principle. The potentiometric arrangements and the setting of the circuit to solve n linear simultaneous equations in n unknowns are explained. Details of the model are given with the help of diagrams.

Modifications to be made in the circuit to extract the eigen values and eigen vectors of a matrix A, when the iterative method of Frost and Tamres is used, are considered. The computer described in this article can handle only a third order secular equations; but the author points out that there is no essential difficulty in constructing a larger computer based on these principles by making use of switches having longer capacity. Improvements in the model and the method of solving the equation $AX=ABX$ will be discussed in Part III (to be published).

(P. K. Kumaran)

MENDENHALL, W. (London School of Economics and North Carolina State Coll.) **11.9** (2.0)
A bibliography on life testing and related topics—*In English*
Biometrika (1958) **45**, 521-543 (approx. 620 refs.)

This select bibliography covers papers concerned with statistical theory and methods applicable to the study of the life characteristic of some biological or physical body.

Although many of the statistical methods developed for life tests of industrial products are applicable to studies of life in other fields, and vice versa, major emphasis is given to the industrial applications.

The special conditions associated with the life test and analysis of its observations are briefly considered : in particular the occurrence of ordered observations from censored samples with or without truncation of the population—and the accelerated test. The introductory matter to this select bibliography notes that very little has been published in connection with this last topic. Other topics of considerable interest in the study of the life characteristic are : extreme value theory, fatigue and wear tests, system reliability and the treatment of sensitivity data. The references have been classified as belonging to one or more of nine groups, as follows :

C. Censored sampling, or sampling from a truncated distribution—both univariate and multivariate.
R. Sampling with renewal.
O. Order statistics. (Excluding papers on extreme value theory or censored sampling.)
E. Extreme value theory.
K. Papers concerned with failure rates and conditional failure density. Also included are tests of randomness for a sequence of events occurring in time.
F. Fatigue testing and wear problems.
M. Machine productivity problems. This group includes some representative papers concerned with a group of machines subject to various failure laws and servicing arrangements.
S. System reliability.
D. Methods applicable to sensitivity data : including the fitting of dosage-mortality curves.

In the bibliography the references are presented in an alphabetical order by authors with the appropriate group classification (indicated by the above capital letters) shown in parentheses at the end of each reference. Some of the groups are obviously subtopics of others but the system of classification has been kept relatively simple, e.g., references under censored sampling are not represented under order statistics. Some of the groups are represented only by some of the more important papers which may be consulted for additional references. In general, the main body of continental literature is represented only by papers revealed by the references given at the end of other papers.

(Wm. R. Buckland)

MOOS, H. **11.5** (6.1)
Calculation of multiple correlations by means of electronic computers—*In German*
Metrika (1958) **1**, 148-153 (1 ref.)

The paper deals with a method of calculating regression coefficients and coefficients of determination by means of an electronic computer (IBM 604). The paper discusses the calculation of the covariance matrix and indicates the number of required products. For the solution of the system of linear equations necessary to obtain the regression coefficients the method of Gauss is used. (The method of Gauss-Doolittle, which is not mentioned by the author, would allow for a remarkable reduction of the computation-work by making use of the symmetry of the covariance matrix.) The paper further indicates a method of calculating the coefficients of determination by means of solving a system of linear equations. The author shows that the amount of computation is smaller than in the case of using direct calculation. For a more detailed discussion of the method of Gauss-Doolittle and the use of punchcard equipment the reader is referred to "Practice of multivariate correlation calculations", *Reihe Mathematische Statistik*, edited by Institut für Praktische Mathematik, Prof. Dr. A. Walther, Technische Hochschule, Darmstadt, Parts I, II and III.

(J. Roppert)

SIOTANI, M. & OZAWA, M. (Institute of Statistical Mathematics, Tokyo) **11.1 (5.4)**

Tables for testing the homogeneity of k independent binomial experiments on certain event based on the range—*In English*

Ann. Inst. Statist. Math., Tokyo (1958) **10**, 47-63 (1 ref., 15 tables)

This paper gives tables for testing the homogeneity of independent binomial experiments. The test criterion is based upon the range of the observed number of occurrences v_i in the k experiments of N trials. In this paper is considered the case where N is not large enough to apply the Normal approximation to the single succession of trials but where, under the null hypothesis : $H: p_1 = p_2 = \ldots p_k \equiv p$ (say), Nk is large enough to regard the estimate $\hat{p} = \Sigma v_i / Nk$ as the true value of p.

Let v_i $(i = 1, 2, \ldots, k)$ be the observed number of occurrences, and p_i $(i = 1, 2, \ldots, k)$ the probability of a certain event in the ith binomial experiment of N trials. The test criterion is defined as : $R_k(N,p) = v_1 - v_2$ where v_1 and v_2 are the largest and smallest values, respectively, in the set of k numbers $v_1, v_2, \ldots, v_k$ and p is the probability of the event under the null hypothesis.

The tables in the paper were computed on an automatic relay computer for $N = 10(1)20, 22, 25, 27$ and 30 ; $k = 2(1)15$; $p = 0{\cdot}50$ $(0{\cdot}10)$ $0{\cdot}10$. The entries in the tables, while they are related to the probabilities

$$P\{R_k(N, p) \geqq r_k\}$$

are, in fact, the greatest value of r_k such that the probability $P\{R_k(N, p) \geqq r_k\}$ does not exceed ten specimen values of the level of significance (α) for fixed N, p and k. The tabulated values of α are $0{\cdot}001$, $0{\cdot}005$, $0{\cdot}01$ $(0{\cdot}01)$ $0{\cdot}06$, $0{\cdot}08$, $0{\cdot}10$.

The reason for adopting this device of tabulation is solely on account of clarity and economy of presentation. Attention is drawn to certain problems of rounding common with tables of this kind and a numerical illustration is given.

(C. Hayashi)

STATISTICAL OFFICE OF THE UNITED NATIONS (Washington) **11.8 (–.–)**

Methods Used in Compiling the United Nations Price Indexes for Basic Commodities in International Trade—*In English*

Statist. Papers ; Series M, No. 29 (1958) (66 pp.)

In the issue for the last month of each quarter of its *Monthly Bulletin of Statistics*, the Statistical Office regularly publishes indexes showing different aspects of the price movements of basic commodities in international trade. Half the weight in certain of these is assigned to export prices (f.o.b.) and half to import prices (c.i.f.) in order to take account of the fact that the price determining markets for a number of primary commodities are in the importing countries and that, for them, export rather than import prices are subject to variation caused by change in freight rates and the fact that, for a number of important currents of trade, import prices are available and export prices are not. The construction of these indexes depends on the methodological innovation of dividing weights evenly within each commodity between export and import prices. This is accomplished by appropriately dividing between the export and import prices the weight of each current of trade for which both export and import prices are available. The manner of making the division is described in general and also specifically for each of the 47 primary commodities entering the indexes in question. The paper furthermore lists the price series entering the indexes, gives the weight of each and sets out, for each commodity, the economic considerations that determined the weighting pattern.

French and Spanish editions are in preparation.

(W. R. Leonard)

VAN YZEREN, J. (Netherlands Central Bureau of Statistics)
A note on the useful properties of Stuvel's index numbers—*In English*
Econometrica (1958) **26**, 429-439 (1 nomogram)

11.8 (11.3)

This paper develops the algebra of the formulae introduced by G. Stuvel [*Econometrica* (1957) **25**, 123-132] for price and quantity index numbers. The principal feature of these indices is the synthesis of the additive and multiplicative way of decomposing value changes. Whereas Laspeyres' indices generally have an upward bias, in the Stuvel indices the biases cancel out. Stuvel's indices, unlike Laspeyres' and Paasche's, satisfy the time reversal test.

The properties of Fisher, Edgeworth and Stuvel indices are considered so far as the aggregation of the two sub-aggregates into one index is concerned. Whereas, with the first two types of index number, the aggregation cannot be carried out using value data only, it is shown that there is a simple and symmetric way of aggregating Stuvel indices with the help of value weights. A further point is that if two Stuvel indices are equal, then the overall index will have the same value, whereas both the Fisher and Edgeworth formulae violate this requirement.

It is shown that Stuvel's indices are unique in that similar indices derived from Paasche, and not Laspeyres' formulae, do not have the same derivable aggregative properties.

Algebraic formulae are then developed in order to compare the values taken by Stuvel indices with those taken by Laspeyres, Paasche, Fisher and Edgeworth and Stuvel indices. Under given conditions, Fisher, Edgeworth and Stuvel indices are intermediate between those of Laspeyres and Paasche.

The position of Edgeworth and Fisher price indices with respect to Stuvel's depends on the absolute value of the Edgeworth index and on the relative values of the corresponding Laspeyres' price and quantity indices. If there is no price change according to Edgeworth then there is none according to Stuvel either.

An empirical test with some numerical data is carried out, which suggests that Fisher and Stuvel indices meet the circular test equally well.

Finally, the proportionality test is discussed, which the Stuvel indices do not satisfy. A way out of this difficulty is suggested which involves the modification of the formulae proposed by Stuvel by the introduction of an additional parameter, which would normally take the value 1, but could, if (say) a monetary reform were carried through, be adjusted to allow for it. The introduction of this extra parameter does not destroy the other desirable properties of the Stuvel indices.

(D. J. Day)

VINSKI, IVO. (Ekonomski Institut, Zagreb)
Problems in Calculating National Wealth—*In German*
Metrika (1958) **1**, 158-172 (13 refs., 3 tables)

11.0 (8.0)

After a historical introduction (I) the article discusses conceptual problems (II), the relationship between national income and national wealth (III), the rate of growth of national wealth (IV), and international comparisons (V).

Since estimates of national wealth were in vogue during the eighteenth and nineteenth centuries no further methodological progress has been made since the earliest thirties. But with the rapid progress of modern national accounting after World War II, national wealth again got attention in the professional world.

National wealth contains natural resources and produced goods. Of the former, only cultivated soil is included since unused natural forces etc. cannot be evaluated. Of the latter, museums, monuments, materials for the armed forces, are excluded for want of an economic evaluation. Differences between the theorists are much greater than between statisticians engaged in those calculations, since the statistical material available in most countries reduces too ambitious concepts. Formerly, economists preferred an indirect method of evaluation, i.e. by capitalisation of income flows. The author regards this method as obsolete and inaccurate. Apart from that it is not applicable if the greater part of wealth is not in private hands. Modern procedures are based on direct evaluations by surveys either in form of complete enumerations (census of agriculture, industry, trade, public utilities, etc.) or in the form of sample surveys (dwellings, furniture, stocks in households, etc.). If such surveys are not simultaneous, the results can be synchronised by Goldsmith's "perpetual inventory method", i.e. by adding net investment to capital stocks of back years. The crucial problem of all national wealth estimates is the principle of evaluation itself. Of the three possibilities: (1) original purchase values, (2) market prices (i.e. sale values) and (3) reproduction values (gross and net of depreciation), the author argues for the latter and presents —with due regard for the unreliability of depreciation—all his figures for Yugoslavia in two series (gross and net).

The relationship of national income to national wealth seems not to be a plain one. It is not simply an indicator of capital productivity, but depends on natural conditions, labour force, terms of trade, etc. Different stages in economic development, therefore, produce relationships of different size. In underdeveloped as well as in highly developed countries the figure is relatively small (national wealth equals about two to three times the national income), in countries with forced industrial development (as Yugoslavia) the figure is about four to five. In a similar way the rate of growth of national wealth depends on the stage of growth already arrived at: it seems to flatten out with the increasing growth and age of the economy.

With regard to international comparisons the author recommends not to rely on *per capita* figures expressed in arbitrary exchange units but to compare the structures of country totals of national wealth. The comparison in the author's table III is convincing.

(L. Bosse)